IDA

INSTITUTE FOR DEFENSE ANALYSES

Comparisons of Transport and Dispersion Model Predictions of the European Tracer Experiment: Area-Based and Population-Based Measures of Effectiveness

Steve Warner, Project Leader
Nathan Platt
James F. Heagy

October 2004

IDA Paper P-3915

Log: H 04-001200

This work was conducted under contract DASW01 04 C 0003, Task DC-9-1797, for the Defense Threat Reduction Agency. The publication of this IDA document does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official position of that Agency.

I N S T I T U T E F O R D E F E N S E A N A L Y S E S

IDA Paper P-3915

Comparisons of Transport and Dispersion Model Predictions of the European Tracer Experiment: Area-Based and Population-Based Measures of Effectiveness

Steve Warner, Project Leader
Nathan Platt
James F. Heagy

PREFACE

This paper was prepared by the Institute for Defense Analyses (IDA) for the Defense Threat Reduction Agency (DTRA), in partial fulfillment of the task "Support for DTRA in the Validation Analysis of Hazardous Material Transport and Dispersion Prediction Models." The objective of this effort was to conduct analyses and special studies associated with the verification, validation, and accreditation (VV&A) of hazardous transport and dispersion prediction models.

This paper is the second in a series of papers that compares the predictions of several transport and dispersion models to the data collected during the *European Tracer Experiment (ETEX)* release of October 1994. The first paper focused on the methodology of comparison – that is, the previously described Measure of Effectiveness (MOE) for transport and dispersion models. This paper extends this effort to consider area-based and population-based MOEs.

The IDA Technical Review Committee was chaired by Robert R. Soule and consisted of Arthur Fries, Vincent B. Lillard, Nelson S. Pacheco, and Edward T. Toton. The authors thank Stefano Galmarini (Joint Research Centre of the European Commission) for providing access to the model predictions of the *ETEX* release and for numerous useful discussions.

COMPARISONS OF TRANSPORT AND DISPERSION MODEL PREDICTIONS OF THE EUROPEAN TRACER EXPERIMENT: AREA-BASED AND POPULATION-BASED MEASURES OF EFFECTIVENESS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SUMMARY

A. INTRODUCTION

In October 1994, the inert, environmentally safe tracer gas perfluoro-methyl-cyclohexane (PMCH) was released over a 12-hour period from a location in northwestern France and tracked at 168 sampling locations in 17 countries across Europe extending over a thousand kilometers.[1] This release, known as the European Tracer Experiment (*ETEX*), resulted in the collection of a wealth of data that can be used to assess long-range transport and dispersion model predictions. IDA has obtained from the Joint Research Centre of the European Commission (Ispra, Italy) 46 sets of transport and dispersion predictions associated with models from 17 countries (Table 1-1) – including HPAC/SCIPUFF and ARAC (LLNL)[2] – as well as the observed PMCH sampling data associated with the October 1994 *ETEX* release.[3]

Recently, a previously developed user-oriented two-dimensional measure of effectiveness (MOE) was used to evaluate the predictions of these 46 models against the long-range *ETEX* observations.[4] The two-dimensional MOE allows for the evaluation of

---

1 Graziani, G., Klug, W., and Mosca, S., 1998: *Real-Time Long-Range Dispersion Model Evaluation of the* ETEX *First Release*, Joint Research Center, European Commission, Office of Official Publications of the European Communities, L-2985 (CL-NA-17754-EN-C), Luxembourg, 1998.

2 HPAC = Hazardous Prediction and Assessment Capability, SCIPUFF = Second-Order Closure Integrated Puff, ARAC = Atmospheric Release Advisory Center, and LLNL = Lawrence Livermore National Laboratory. HPAC/SCIPUFF and ARAC (now known as NARAC – National ARAC) are of particular interest to our sponsor.

3 Mosca, S., Bianconi, R., Bellasio, R., Graziani, G., and Klug, W., 1998: *ATMES II – Evaluation of Long-Range Dispersion Models Using Data of the 1st* ETEX *Release*, Joint Research Centre, European Commission, Office of Official Publications of the European Communities, L-2985 (CL-NA-17756-EN-C), Luxembourg, 1998.

4 Warner, S., Platt, N., and Heagy, J. F., 2003: *Application of User-Oriented MOE to Transport and Dispersion Model Predictions of the European Tracer Experiment*, IDA Paper P-3829, November 2003. (Available electronically [DTIC STINET ADA419433] or on CD via e-mail request to Steve Warner at swarner@ida.org or a mail request to Steve Warner, Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, Virginia 22311-1882.); Platt, N., Warner, S., and Heagy, J. F., 2004: "Application of User-Oriented MOE to Transport and Dispersion Model Predictions of ETEX," *Proceedings of the Ninth International Conference on Harmonisation Within Atmospheric Dispersion Modelling for Regulatory Purposes*, Garmisch-Partenkirchen, Germany, 1-4 June 2004, pages 120-125; and Warner, S., Platt, N., and Heagy, J. F., 2004: Application of user-oriented measure of

transport and dispersion model predictions in terms of "false negative" (under-prediction) and "false positive" (over-prediction) regions.[5] A perfect model prediction leads to no false negative *and* no false positive, that is, complete and perfect overlap of the predictions and observations. Such a perfect model would have a two-dimensional MOE value of (1,1).[6] For a given application and user risk tolerance, certain regions of the two-dimensional MOE space may be considered acceptable. For example, some users may tolerate a certain false positive fraction (ultimately, unnecessarily warned individuals) but require a very low false negative fraction (inadvertently exposed individuals). Such a risk tolerance profile implies a certain location in the two-dimensional MOE space, which can be turned into a mathematical function for "scoring" the MOE predictions. User "scoring" functions have been previously developed for the MOE.[7]

The MOE values were previously computed by considering the prediction of concentrations summed across all sampler locations and based on defining a critical threshold. For threshold-based MOE values, the model was judged by its ability to predict which locations led to observations above certain specified thresholds. Threshold-based MOE values were calculated for three concentrations:[8] 0.01, 0.1, and 0.5 ng m^{-3}. *The current report extends this previous work by computing MOE values that*

---

effectiveness to transport and dispersion model predictions of the European tracer experiment. *Atmos. Environ.*, in press.

5 Warner, S., Platt, N., and Heagy, J. F., 2004: "User-Oriented Two-Dimensional Measure of Effectiveness for the Evaluation of Transport and Dispersion Models," *J. Appl. Meteor.,* **43**: 58-73 and Warner S., Platt, N., and Heagy, 2001: "User-Oriented Measures of Effectiveness for the Evaluation of Transport and Dispersion Models," *Proceedings of the Seventh International Conference on Harmonisation Within Atmospheric Dispersion Modelling for Regulatory Purposes*, Belgirate, Italy, 28-31 May 2001, pages 24-29.

6 A model prediction that completely misses the observation (perhaps, the "plume" goes in the exact opposite direction) would achieve an MOE value of (0,0).

7 Warner, S., Platt, N., and Heagy, J. F., 2001: *Application of User-Oriented MOE to HPAC Probabilistic Predictions of Prairie Grass Field Trials*, IDA Paper P-3586, May 2001. (Available electronically [DTIC STINET ada391653] or on CD via e-mail request to Steve Warner at swarner@ida.org or a mail request to Steve Warner, Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, Virginia 22311-1882.)

8 The threshold levels of 0.01, 0.1, and 0.5 ng m^{-3} correspond *exactly* to previously examined values. See Mosca, S., Graziani, G., Klug, W., Bellasio, R., and Bianconi, R., 1998: A statistical methodology for the evaluation of long-range dispersion models: an application to the ETEX exercise. *Atmos. Environ.,* 32 (24), 4307-4324 and Boybeyi, Z., Ahmad, N., Bacon, D. P., Dunn, T. J., Hall, M. S., Lee, P. C. S., Sarma, R. A., and Wait, T. R., 2001. Evaluation of the Operational Multiscale Environment model with Grid Adaptivity (OMEGA) against the European Tracer Experiment (ETEX). *J. Appl. Meteor.,* **40**, 1541-1558. Also, the experimenters considered the value 0.01 ng m^{-3} has a lower bound for examinations. See the references cited in footnote 4 for additional details.

are based on "true" areas (e.g., in square kilometers) and on actual European population distributions. The overall objective of this paper is to document the procedures used to estimate the area-based and population-based MOE values. In addition, studies of the sensitivity of the relative rankings of model predictions to the differing computational techniques are discussed.

B. RESULTS AND DISCUSSION: MODEL COMPARISONS TO *ETEX*

1. Area- and Population-Based MOEs

a. Area-Based

In order to create *area-based* MOE values, two procedures were developed for this study. First, a procedure was used that assigns weights to each sampler location that corresponds to an assessment of the area *represented* by each sampler. Basically, samplers placed closer together account for less area than samplers placed further apart. Figure 1 shows the locations of the 168 PMCH samplers across the European continent (red triangles) as well as the release location in northwestern France (black circle). A diagram (referred to as a Voronoi diagram) can be created by dividing Europe into regions (polygons) such that each region corresponds to one of the sampler locations and all of the points in a given region are closer to the corresponding sampler location than to any other sampler location. Given this diagram, each sampler observation/prediction comparison can be weighted by the area of the corresponding Voronoi polygon. Area-based MOE values can then be created directly from these weighted observation/prediction comparisons. The details of this technique are described in Chapter 1.

In addition to the above weighting technique, area-based MOE values were created using interpolation to create intermediate values on a regular grid from the values observed (or predicted) at the discrete sampler locations. For this procedure, we first transformed the data (observations or predictions) logarithmically and then followed a Delaunay triangulation procedure with linear interpolation. The details of this technique are described in Chapter 1.

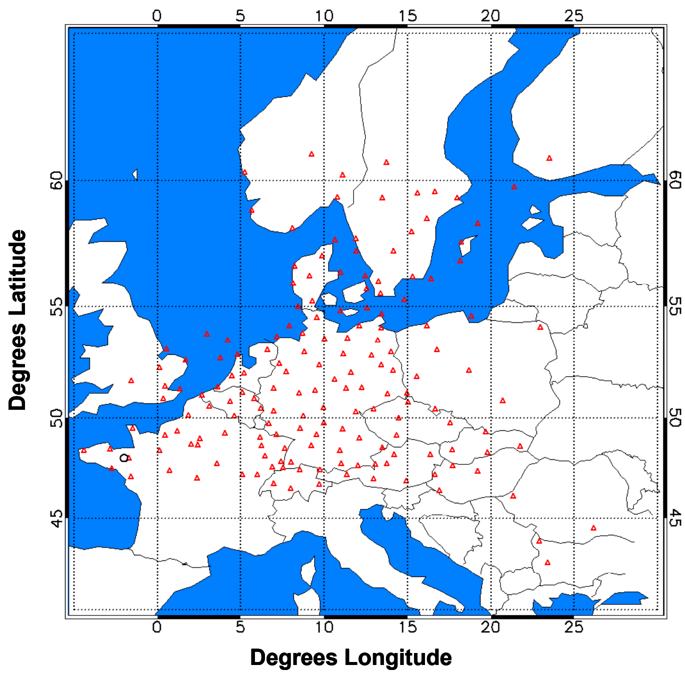


Figure 1. ***ETEX* Sampler Locations Across Europe: Red Open Triangles Correspond to Sampler Locations and the Black Open Circle Corresponds to the Release Location.**

b. Population-Based

In order to create population-based MOE values, two extensions of the previously described area-based MOE techniques are required. First, dosages must be created from the observed and predicted concentrations. This can be accomplished by summing concentrations from distinct 3-hour time periods at each of the sampler locations.[9] For this analysis, three threshold dosages were examined: 7.2, 72, and 360 ng min m^{-3}. These three values can be considered as related to the 3-hour average concentration thresholds of 0.01, 0.1, and 0.5 ng m^{-3} by considering a 12-hour (720 min) period in which the cloud might pass over any individual sampler location.[10]

Next, the underlying non-uniform European population distribution must be taken into account in order to convert to population-based MOE values. Figure 2 illustrates the population distribution that was used. The population distribution shown is represented

---

9 To create "observed" dosages at given locations, one simply sums the concentrations at each sampler location. However, periods of time in which sampler data could not be (or were not) collected exist for many of the sampler locations. For these time periods, spatial interpolation provides a natural way to fill in the temporal holes in the observed concentration data and was used for some of the area-based dosage calculations.

10 The release duration was 12 hours.

by population values at about 2.1 million grid cells with a grid cell resolution of about 2 km by 2 km. The overall European population represented here is about 500 million. At this point then, for a given threshold, MOE values can be expressed in terms of the "fraction of the population inadvertently exposed" and the "fraction of the population unnecessarily warned" – i.e., population-based MOE values.

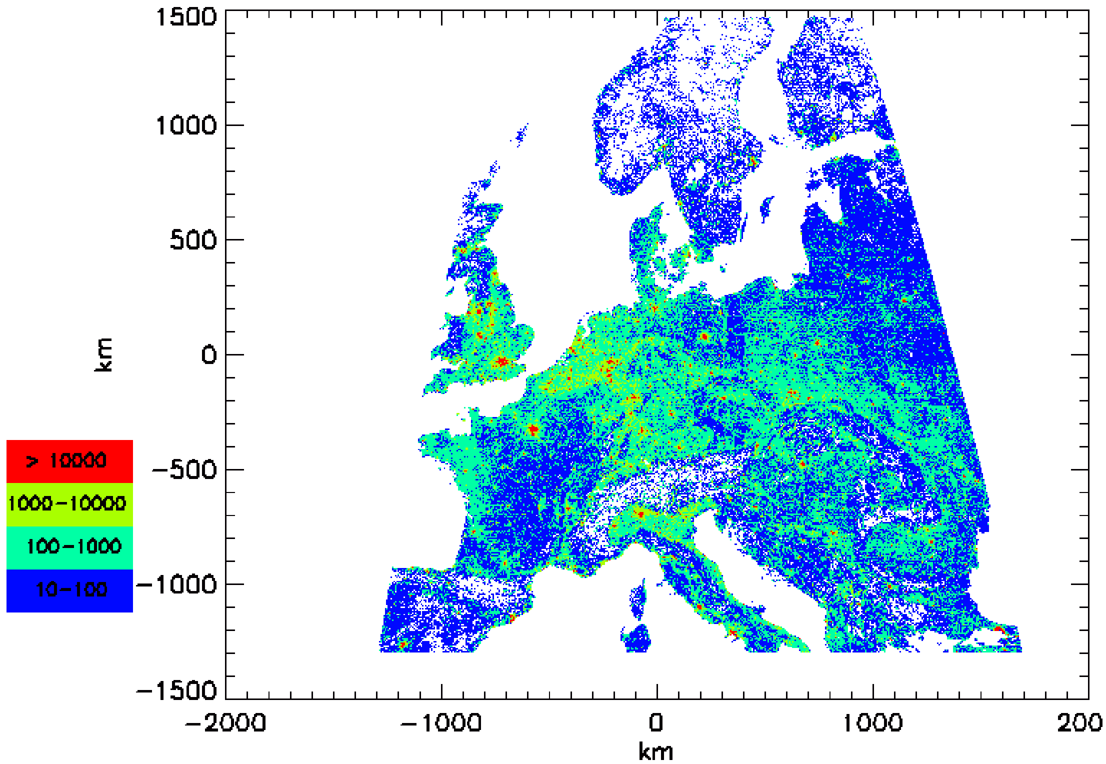


Figure 2. Illustration of European Population Distribution

2. Comparisons of Model Rankings: Area- and Population-Based MOE Values

One goal of this study was to develop insight into how differences in techniques for computing the MOE might change conclusions with respect to model performance. We focused on relative model behavior and used various scoring functions to rank model performance. Then, changes in model rankings were used as the basis to evaluate the sensitivity of assessments of relative model performance to differing computational techniques. In particular, we examined the following two sets of metrics. First, when comparing two MOE computational techniques to each other, the mean and median of the absolute value of the ranking differences for the 46 sets of model predictions were examined. For example, a mean value in absolute ranking difference of 3 implies that on average (for the 46), models changed their relative ranked position by 3. Histograms of ranking differences were also examined. Next, we were interested in the robustness of the differing area-based and population-based MOE techniques at detecting the "best" and "worst" performing models. As such, the fraction of models ranked in the top ten (or bottom ten) for any two techniques being compared was also examined.

Based on the above studies, the biggest causes of variance in terms of changes to relative model rankings could be identified. First, it must be noted that, overall, relative model rankings were reasonably robust to the computational techniques that were examined. For instance, it was typical for 7 or more of the top (or bottom) 10 ranked models to remain in the top (or bottom) 10 when comparing two techniques. The biggest source of variance in relative model rankings was associated with comparisons of concentration and dosage MOE values. Area-based MOE values were computed for 3-hour average concentration thresholds and for 12-hour dosage thresholds. A few models improved their relative rankings greatly when assessed on dosage MOE values instead of concentration-based values. For example, based on the MOE value closest to (1,1) and for the three comparative concentration/dosage thresholds (0.01 ng m^{-3} / 7.2 ng min m^{-3}, 0.1 ng m^{-3} / 72 ng min m^{-3}, 0.5 ng m^{-3} / 360 ng min m^{-3}), model 121 (SCIPUFF) moves up 19 (from 21 to 2), 12 (from 23 to 11), and 8 (from 29 to 21) positions, respectively. Examination of 3-hour average concentration plots of the observations and predictions suggests that some models do not match the concentration timing (e.g., "plume" arrival and dwell time) as well as others. Therefore, while dosages may be well predicted, concentrations, that require both the location *and* time to be matched, may be predicted worse (relative to the other models). Chapters 2 and 3 of this paper also describe smaller changes in relative model rankings associated with changes in the computational techniques used to estimate the area size (sampler weighting versus interpolation) and the inclusion of actual European populations.

Next, it was found that some sets of model predictions led to top 10 model performance for many or even all of the MOE computational techniques. These "robust" predictions are highlighted in the main body of this paper, and included at or near the top of the list were the Canadian Meteorological Center's *ETEX* predictions (model 105) and the Lawrence Livermore National Laboratory's *ETEX* predictions (model 127).

The rankings described in this paper result from consideration of a single release and *general* inference about which model is "best" or ranked highest is not appropriate. Rather, these rankings describe performance in terms of this specific release only. In addition, for this single release field experiment, no direct measures of uncertainty associated with the computed MOE values or model rankings were constructed. Previous studies that have examined multiple releases have described techniques for assessing uncertainties and comparing metrics to identify statistically significant differences.[11]

Finally, an important caveat must be noted. To this point, the use of area-based and population-based MOE values to compare sets of model predictions, rank the models, and provide insight into relative model performance has been emphasized. With respect to the population-based two-dimensional (i.e., x and y axes) MOE values, the x-axis corresponds to one minus the fraction of the (exposed) population that is inadvertently exposed (i.e., "not warned") to a threshold level of interest and the y-axis corresponds to one minus the fraction of the (warned) population that is unnecessarily warned (at a threshold level of interest). One might imagine using an effects (or lethality) model to compute, via minimal extension of the MOE, the actual number of people "falsely warned" or "inadvertently exposed." However, one must be careful because of the relatively small number of samplers associated with the observed *ETEX* data. In attempting to describe the actual number of affected people, one would need to rely on the absolute (actual) areas computed, not simply the fraction of areas. In such a case, the estimated area sizes can be sensitive to the details associated with the specific area-based technique (e.g., interpolation) used.

C. OUTLINE OF THIS PAPER

This paper is divided into three chapters. Chapter 1 briefly describes the *ETEX*, reviews the user-oriented MOE and associated scoring functions, and summarizes previous results and discussions associated with comparisons of the 46 model predictions. Chapter 2 describes two methodologies – sampler weighting and interpolation – for creating area-based MOE values associated with concentration predictions and observations. Model rankings associated with area-based MOE values and specific scoring functions are described in this chapter. In addition, Chapter 2 reports comparisons of area-based and nominal (those computed in the previous report)[12] MOE values and subsequent model rankings as well as comparisons between sampler-weighted and interpolated area-based MOE values and associated model rankings. Chapter 3 starts with a description of the computation of dosage-based MOE values for *ETEX* predictions. Then, using the area-based methodologies of Chapter 2 and the actual European

[11] For example, see Warner, S., Platt, N., and Heagy, J. F., 2004: Comparison of transport and dispersion model predictions of the *Urban 2000* field experiment. *J. Appl. Meteor.*, **43**: 829-846.

[12] See the references cited in footnote 4.

population distribution, MOE values characterized by an x-axis labeled “one minus the fraction of the population inadvertently exposed” and a y-axis labeled “one minus the fraction of the population unnecessarily warned” are created. These population-based MOE values are then used, together with scoring functions to rank the 46 models. Appendix A provides a list of acronyms and Appendix B provides technical details associated with a coordinate system conversion that was required during this analysis. Appendix C provides an extract from the task order that supported this research.

CHAPTER 1

INTRODUCTION

1. INTRODUCTION

A. THE 1994 EUROPEAN TRACER EXPERIMENT (*ETEX*)

In October 1994, the inert, environmentally safe, tracer gas perfluoro-methyl-cyclohexane (PMCH) was released over a 12-hour period from a location in northwestern France and tracked at 168 sampling locations in 17 countries across Europe [Refs. 1-1 through 1-3]. This release, known as *ETEX*, resulted in the collection of a wealth of data. IDA has obtained from the Joint Research Centre of the European Commission (Ispra, Italy), 46 sets of transport and dispersion predictions associated with models from 17 countries – including HPAC/SCIPUFF and ARAC (LLNL)[1] – as well as the observed PMCH sampling data associated with the *ETEX* release.

The *ETEX* release began at 16:00 UTC[2] on 23 October 1994 and ended at 3:50 UTC on 24 October 1994. The release location was 35 km west of Rennes (Monterfil, 20°00'20"W, 48°03'30"N) in Brittany, France. PMCH was released 8 m above ground level at a rate of 7.95 g s^{-1}.

The samplers were located at synoptic stations of the various national meteorological services. Air samples were collected for 3 hours, every 3 hours, for a period of 90 hours after the initial release. Figure 1-1 shows the locations of the samplers across Europe.

Measurements of PMCH were made before, during, and after the release at several stations and average background levels were subtracted from the measured data. Furthermore, these measurements suggested that a level of 0.01 ng m^{-3} should be used as the minimum for all statistical comparisons.

A detailed description of the weather situation during the experiment (23-26 October 2004) can be found in Ref. 1-1 (pages 44 – 50) including the synoptic charts for

---

1 HPAC = Hazardous Prediction and Assessment Capability, SCIPUFF = Second-Order Closure Integrated Puff, ARAC = Atmospheric Release Advisory Center, and LLNL = Lawrence Livermore National Laboratory. HPAC/SCIPUFF and ARAC (now known as NARAC – National ARAC) are of particular interest to our sponsor.

2 UTC = Universal Time Coordinated.

each day. Winds were generally from the west, and then later from the southwest, and the continent was dominated by a low pressure zone centered east of Scotland in the North Sea.

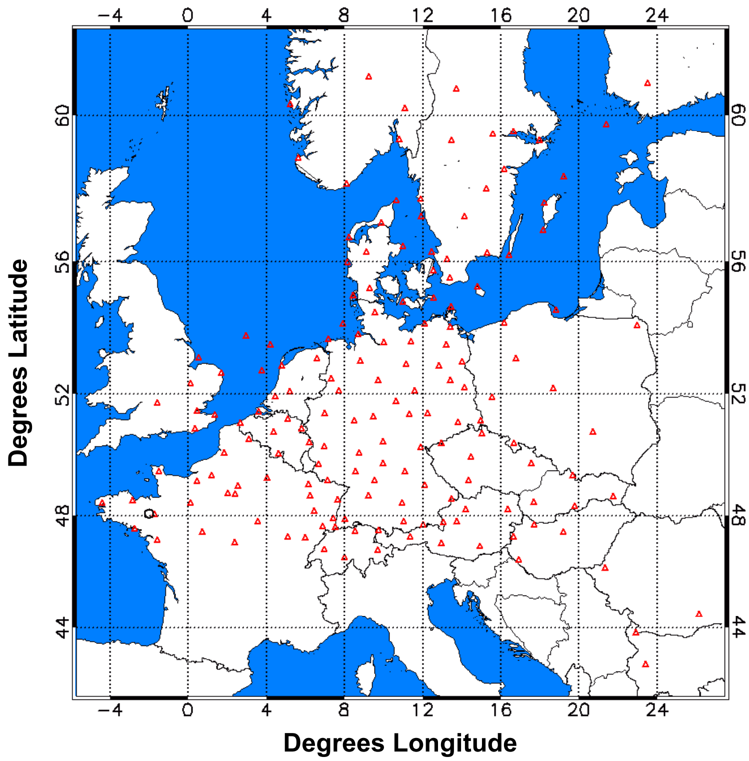


Figure 1-1. *ETEX* Sampler Locations Across Europe: Red Open Triangles Correspond to Sampler Locations and the Black Open Circle Corresponds to the Release Location.

Two years after the *ETEX* releases, a modeling exercise known as ATMES II was conducted.[3] *ETEX*-ATMES II predictions associated with 46 model configurations were provided to IDA by the Joint Research Centre of the European Commission (Ispra, Italy).[4] Table 1-1 provides some details associated with these models.[5] The series of model predictions denoted with numbers between 101 and 135 (the "100 series") used European Centre for Medium Range Weather Forecasts (ECMWF) analyzed meteorological data as input. The "200 series" (201-214) used weather inputs selected by the modeler and not the ECMWF-related data. Comparisons between 100 series

3 ATMES = Atmospheric Transport Model Evaluation Study.

4 These predictions can be downloaded from the *ETEX* public access web sites: http://rem.jrc.cec.eu.int/atmes2/ and http://rem.jrc.cec.eu.int/etex/.

5 An additional three sets of predictions associated with the Royal Dutch Meteorological Institute were not available to us but were part of the original *ETEX* (ATMES II) study [Ref. 1-2]. Table 1-1 is extracted from Reference 1-2.

predictions were posited to identify differences related to variations in dispersion modeling. Comparisons between 100 and 200 series predictions of the same underlying transport and dispersion model should emphasize differences associated with the input wind field. In Table 1-1, model 121 (DTRA's SCIPUFF) and model 127 (LLNL's ARAC) are highlighted in red bold.

Table 1-1. ATMES II Participants for which IDA Obtained Predictions

Model	Acronym	Participant	Nationality
101	IMP	Institute of Meteorology and Physics, University of Wien	Austria
102	BMRC	Bureau of Meteorology Research Centre	Australia
103	NIMH-BG	National Institute of Meteorology and Hydrology	Bulgaria
104	NIMH-BG	National Institute of Meteorology and Hydrology	Bulgaria
105	CMC	Canadian Meteorology Centre	Canada
106	DWD	German Weather Service	Germany
107	DWD	German Weather Service	Germany
108	NERI	Nat. Environment Research Inst./Risoe Nat. Lab./Univ. of Cologne	Germany/Denmark
109	NERI	Nat. Environment Research Inst./Risoe Nat. Lab./Univ. of Cologne	Germany/Denmark
110	DMI	Danish Meteorological Institute	Denmark
111	IPSN	French Institute for Nuclear Protection and Safety	France
112	EDF	French Electricity	France
113	ANPA	National Agency for Environment	Italy
114	CNR	National Research Council	Italy
115	JAERI	Japan Atomic Research Institute	Japan
116	MRI	Meteorological Research Institute	Japan
117	NIMH-R	National Institute of Meteorology and Hydrology	Romania
118	FOA	Defense Research Establishment	Sweden
119	MetOff	Meteorological Office	United Kingdom
120	NOAA	National Oceanic and Atmospheric Administration	United States
121	ARAP **(SCIPUFF)**	ARAP Group of Titan Research and Technology	United States
122	KMI	Royal Institute of Meteorology of Belgium	Belgium
123	Meteo	Meteo France	France
127	LLNL **(ARAC)**	Lawrence Livermore National Laboratories	United States
128	SMHI	Swedish Meteorological and Hydrological Institute	Sweden
129	SAIC	Science Applications International Corporation	United States
130	IMS	Swiss Meteorological Institute	Switzerland

Table 1-1. ATMES II Participants for which IDA Obtained Predictions (continued)

Model	Acronym	Participant	Nationality
131	DNMI	Norwegian Meteorological Institute	Norway
132	SRS	Westinghouse Savannah River Laboratory	United States
133	JMA	Japan Meteorological Agency	Japan
134	JMA	Japan Meteorological Agency	Japan
135	MSC-E	Meteorological Synthesizing Centre - East	Russia
201	BMRC	Bureau of Meteorology Research Centre	Australia
202	CMC	Canadian Meteorological Centre	Canada
203	DWD	German Weather Service	Germany
204	NERI	Nat. Environment Research Inst./Risoe Nat. Lab./Univ. of Cologne	Germany/Denmark
205	DMI	Danish Metrological Institute	Denmark
206	Meteo	Meteo France	France
207	MRI	Meteorological Research Institute	Japan
208	SMHI	Swedish Meteorological and Hydrological Institute	Sweden
209	MetOff	Meteorological Office	United Kingdom
210	MetOff	Meteorological Office	United Kingdom
211	NOAA	National Oceanic and Atmospheric Administration	United States
212	NIMH-R	National Institute of Meteorology and Hydrology	Romania
213	DNMI	Norwegian Meteorological Institute	Norway
214	MSC-E	Meteorological Synthesizing Centre - East	Russia

B. PREVIOUS RESEARCH: APPLICATION OF USER-ORIENTED MOE TO TRANSPORT AND DISPERSION MODEL PREDICTIONS OF *ETEX*

Recently, user-oriented Measure of Effectiveness (MOE) values for the 46 ATMES II-related transport and dispersion model predictions of *ETEX* have been computed and analyzed [Ref. 1-4]. The results of these recent analyses are briefly described in this section.

1. User-Oriented MOE

In general, model validation efforts include identifying specific metrics that are needed to compare field trial observations and predictions. It is helpful if model validation includes an MOE that relates "operational" use of the model to field trial experiments. Such an MOE gives a certain degree of confidence to users with respect to how closely the model approximates the real world in their particular situation.

Previously, we developed and described a two-dimensional user-oriented MOE [Refs. 1-5 and 1-6] and described several [Refs. 1-7 through 1-15] applications. The two-dimensional MOE allows for the evaluation of transport and dispersion model predictions in terms of “false negative” (under-prediction) and “false positive” (over-prediction) regions. Figure 1-2 shows one possible interpretation of these regions – the observed and predicted areas in which a prescribed dosage is exceeded. This view can be extended to consider the marginal over- and under-predicted values. In any case, numerical estimates of the false negative region (A_{FN}), the false positive region (A_{FP}), and the overlap region (A_{OV}) characterize this conceptual view.

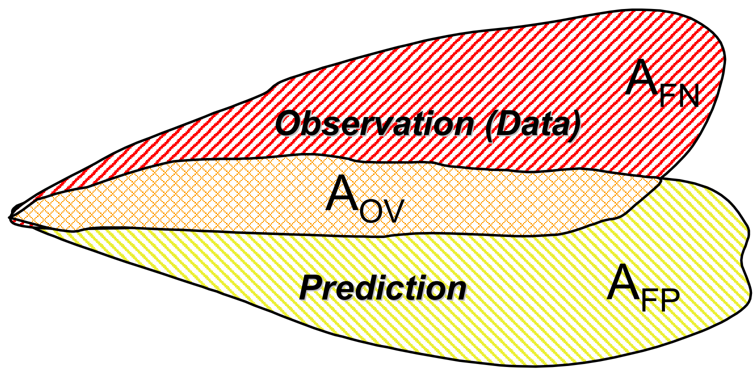


Figure 1-2. Conceptual View of Overlap (A_{OV}), False Negative (A_{FN}), and False Positive (A_{FP}) Regions that Are Used to Construct the User-Oriented MOE

The MOE that we consider has two dimensions. The x-axis corresponds to the ratio of overlap region to the observed region and the y-axis corresponds to the ratio of overlap region to the predicted region. When these mathematical definitions are algebraically rearranged (Eq. 1-1 below), we recognize that the x-axis corresponds to *1 minus the false negative fraction* and the y-axis corresponds *to 1 minus the false positive fraction*.

$$MOE = (x, y) = \left(\frac{A_{OV}}{A_{OB}}, \frac{A_{OV}}{A_{PR}} \right) = \left(\frac{A_{OB} - A_{FN}}{A_{OB}}, \frac{A_{PR} - A_{FP}}{A_{PR}} \right) = \left(1 - \frac{A_{FN}}{A_{OB}}, 1 - \frac{A_{FP}}{A_{PR}} \right). \qquad (1\text{-}1)$$

where A_{FN} = region of false negative, A_{FP} = region of false positive, A_{OV} = region of overlap, A_{PR} = region of the prediction, and A_{OB} = region of the observation. Consistent with the above algebraic rearrangement, Figure 1-3 shows the region of false negative decreasing from left to right and the region of the false positive decreasing from bottom to top.

A perfect model prediction leads to no false negative *and* no false positive, that is, complete and perfect overlap of the predictions and observations. Such a perfect model would have a two-dimensional MOE value of (1,1) as illustrated in Figure 1-3.[6] Additional discussion of the key characteristics of the two-dimensional MOE space can be found in Reference 1-4.

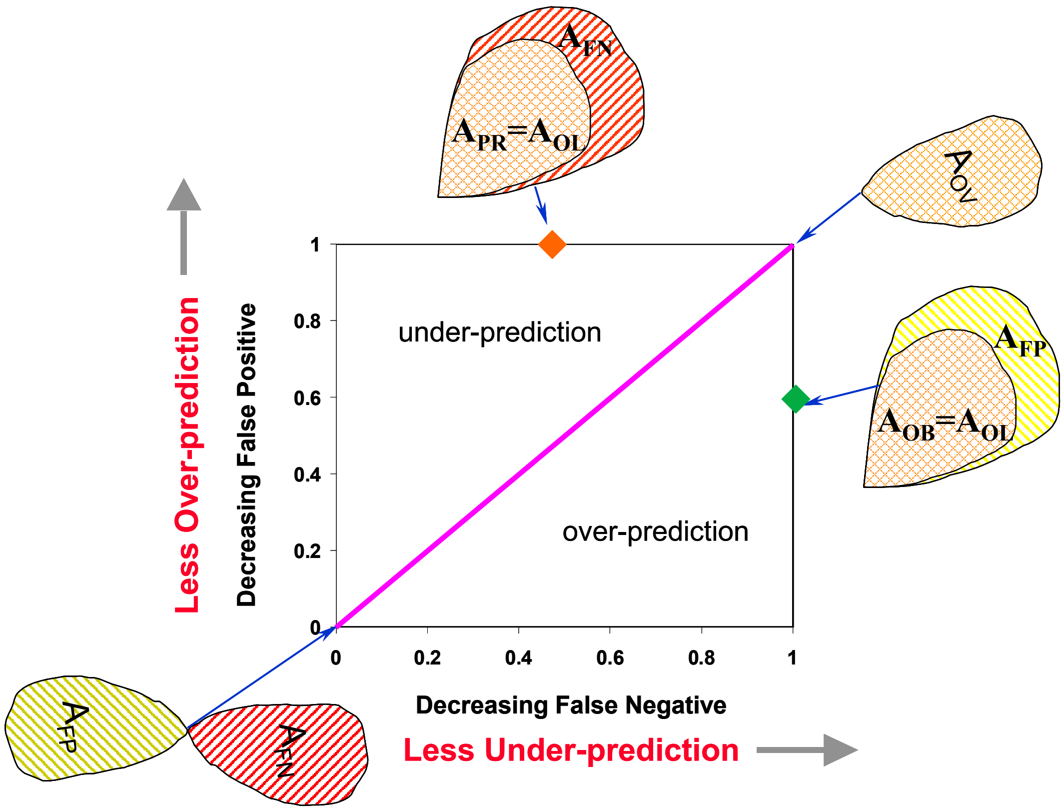


Figure 1-3. Key Characteristics of the Two-Dimensional MOE Space

MOE values can be computed by considering the prediction of concentrations summed across all sampler locations or MOE values can be computed based on defining a critical threshold. For threshold-based MOE values, the model is judged by its ability to predict which locations led to observations above a certain specified threshold. Previously, threshold-based MOE values for predictions of *ETEX* were computed for three thresholds: 0.01,[7] 0.1, and 0.5 ng m^{-3}. Details of these computations are described in Reference 1-4.

6 A model prediction that completely misses the observation (perhaps, the "plume" goes in the exact opposite direction) would achieve an MOE value of (0,0).

7 The value 0.01 ng m^{-3} was considered a lower bound by the experimenters. The experimenters treated any measurement below 0.01 ng m^{-3} as a zero.

2. MOE Scoring Functions and Rankings

For a given application and user risk tolerance, certain regions of the two-dimensional MOE space may be considered acceptable. For example, some users may tolerate a certain false positive fraction (ultimately, unnecessarily warned individuals) but require a very low false negative fraction (inadvertently exposed individuals). Such a risk tolerance profile implies a certain location in the two-dimensional MOE space, which can be turned into a mathematical function for "scoring" the MOE predictions.

The development of notional risk scoring functions is described in Reference 1-4. For threshold-based MOE values, two scoring functions were identified – the objective scoring function (OSF) and the risk-weighted figure of merit in space (RWFMS).[8] The OSF measures the "distance" in the two-dimensional MOE space from the point (1,1). Therefore, the smaller OSF values imply MOE estimates that are closer to the perfect (1,1). The computed OSF values can be used to rank model predictions. OSF is defined as:

$$OSF = d_{OSF} = \sqrt{\left(1 - \frac{A_{OV}}{A_{OB}}\right)^2 + \left(1 - \frac{A_{OV}}{A_{PR}}\right)^2} = \sqrt{\left(\frac{A_{FN}}{A_{OB}}\right)^2 + \left(\frac{A_{FP}}{A_{PR}}\right)^2} \,. \qquad (1\text{-}2)$$

The figure of merit in space (FMS) is defined as the ratio of the intersection of the observed and predicted areas to the union of the observed and predicted areas at a fixed time and above a defined threshold concentration. FMS has been previously defined and used by Mosca, et al. [1998, Ref. 1-16]. RWFMS allows a user to weight the relative contributions of false positives and false negatives within their scoring function as shown below.

$$RWFMS = \frac{A_{OV}}{A_{OV} + C_{FN} A_{FN} + C_{FP} A_{FP}} \qquad (1\text{-}3)$$

where the coefficients, C_{FN} and C_{FP}, are used to weight the false negative and false positive regions, respectively, and C_{FN}, $C_{FP} > 0$. FMS is a special case of RWFMS obtained by setting $C_{FN} = C_{FP} = 1$. As was the case with OSF, RWFMS can be used to rank model predictions. However, in the case of RWFMS, the coefficients C_{FN} and C_{FP} must be stipulated (by the user).

---

8 Other scoring functions for summed concentration-based MOEs were also described in Reference 1-4.

3. Brief Summary of Results and Discussion for Nominal MOE Analyses of *ETEX* ATMES II Predictions [Ref. 1-4]

The user-oriented two-dimensional MOE was used to evaluate 46 sets of model predictions of *ETEX*. Using a few scoring functions that could be identified with notional user requirements, these 46 models were ranked in terms of the desired performance as specified by the scoring function. The sensitivity of MOE values to any single sampler location was examined and it was found that the evaluations of a few models' performance was greatly affected by a single sampler location close to the release point. Finally, the usage of the MOE to explore the time-dependence of model performance was briefly introduced and described.

Table 1-2 identifies the top ranked model predictions as judged by the OSF as well as the rankings (out of 46) of SCIPUFF and ARAC. Rankings are identified for the three threshold-based and summed concentration-based MOE values. No single model dominated the top ranking. Complete rankings can be found in Reference 1-4. The bolded text rankings associated with the SCIPUFF and ARAC summed concentration column, correspond to the rankings of these two sets of model predictions after the removal from consideration of a single close-in sampler located near Rennes, France. While most of the models' MOE values were relatively unaffected by the removal of this single sampler location, a few sets of predictions were, perhaps, overly influenced by this single sampler location. Reference 1-4 discusses this result in more detail.

Table 1-2. Top-Ranked Model and Rankings of SCIPUFF and ARAC Based on MOE Values and the Objective Scoring Function

Rank	0.01 ng m^{-3}	0.1 ng m^{-3}	0.5 ng m^{-3}	Summed Concentration
1	Canadian Meteorological Centre	Swedish Meteorological and Hydrological Office	ARAC	German Weather Service
Model	**0.01 ng m^{-3}**	**0.1 ng m^{-3}**	**0.5 ng m^{-3}**	**Summed Concentration**
SCIPUFF	24	30	23	41 / **34**
ARAC	4	5	1	33 / **8**

When judging model predictive performance using the MOE based on the 0.01 or 0.1 ng m^{-3} threshold, one of two time-dependent behaviors was typically observed in the previous study. For some models, an initial under-prediction of the number of locations that exceed the threshold is followed by a "correction" that leads to about the right number of locations predicted above the threshold, followed finally, by degradation that suggests a general missing of the locations at which the threshold is exceeded at the longest times (and distances). For other models, an initial over-prediction of the number of locations that exceed the threshold is followed by a "correction" that leads to about the right number of locations predicted above the threshold, followed again, by degradation that suggests a general missing of the locations at which the threshold is exceeded.

C. MOTIVATION FOR THIS STUDY

This study extends the previous set of nominal MOE computations and comparisons by considering actual area-based and population-based MOE values. First, MOE values based on actual areas (e.g., square kilometers) must be created. The previous study compared the observations and predictions directly; that is, no area interpolation or sampler weighting was used. An important part of this next effort will be to explore and understand potential sensitivities associated with interpolation/weighting given the underlying sampler space across Europe. Given area-based MOE values, we then consider including the actual European population distributions and notional effects levels of interest to place the MOE in its ultimate context – fraction of the population falsely warned and fraction of the population inadvertently exposed. At this point the 46 models can be re-ranked given this more operational context.

D. OUTLINE OF THIS STUDY

Chapter 2 describes the procedures used to compute area-based MOE values for predictions of *ETEX*. A few different methodologies are described and the resulting area-based-MOE values and rankings are compared with the previous nominal MOE values and rankings. Chapter 3 builds on Chapter 2 and describes the computation of European population-based MOE values and model rankings.

REFERENCES

1-1. Graziani, G., Klug, W., and Mosca, S., 1998: *Real-Time Long-Range Dispersion Model Evaluation of the* ETEX *First Release*, Joint Research Center, European Commission, Office of Official Publications of the European Communities, L-2985 (CL-NA-17754-EN-C), Luxembourg, 1998.

1-2. Mosca, S., Bianconi, R., Bellasio, R., Graziani, G., and Klug, W., 1998: *ATMES II – Evaluation of Long-Range Dispersion Models Using Data of the 1st* ETEX *Release*, Joint Research Center, European Commission, Office of Official Publications of the European Communities, L-2985 (CL-NA-17756-EN-C), Luxembourg, 1998.

1-3. Graziani, G., Galmarini, S., Grippa, G., and Klug, W., 1998: *Real-Time Long-Range Dispersion Model Evaluation of the* ETEX *Second Release*, Joint Research Center, European Commission, Office of Official Publications of the European Communities, L-2985 (CL-NA-17755-EN-C), Luxembourg, 1998.

1-4. Warner S., Platt, N., Heagy, J. F., 2003: *Application of User-Oriented MOE to Transport and Dispersion Model Predictions of the European Tracer Experiment*, IDA Paper P-3829, 86 pp, November 2003. (Available electronically [DTIC STINET ada419433] or on CD via e-mail request to Steve Warner at swarner@ida.org or a mail request to Steve Warner, Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, Virginia 22311-1882); Platt, N., Warner, S., and Heagy, J. F. 2004: Application of user-oriented MOE to transport and dispersion model predictions of ETEX. *Proceedings of the Ninth International Conference on Harmonisation Within Atmospheric Dispersion Modelling for Regulatory Purposes*, Garmisch-Partenkirchen, Germany, 1-4 June 2004, pages 120-125; and Warner S., Platt, N., Heagy, J. F., 2004: Application of user-oriented MOE to transport and dispersion model predictions of the European tracer experiment. *Atmospheric Environment*, in press.

1-5. Warner, S., Platt, N., and Heagy, J. F., 2004: User-oriented two-dimensional measure of effectiveness for the evaluation of transport and dispersion Models," *J. Appl. Meteor.*, **43**: 58-73.

1-6. Warner S., Platt, N., and Heagy, 2001: User-oriented measures of effectiveness for the evaluation of transport and dispersion models. *Proceedings of the Seventh International Conference on Harmonisation Within Atmospheric Dispersion Modelling for Regulatory Purposes*, Belgirate, Italy, 28-21 May 2001, pages 24-29.

1-7. Warner S., Platt, N., Heagy, J. F., Bradley, S., Bieberbach, G., Sugiyama, G., Nasstrom, J. S., Foster, K. T, and Larson, D., 2001: *User-Oriented Measures of Effectiveness for the Evaluation of Transport and Dispersion Models*, IDA Paper P-3554, 797 pp, January 2001. (Available electronically [DTIC STINET ada387239] or on CD via e-mail request to Steve Warner at swarner@ida.org or a mail request to Steve Warner, Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, Virginia 22311-1882.)

1-8. Warner, S., Platt, N., and Heagy, J. F., 2002: *Explorations to Support the Selection of Trials for HPAC/NARAC Model Comparisons: User-Oriented Measure of Effectiveness Values for the Over-Land Along-Wind Dispersion (OLAD) Field Experiments of September 1997*, IDA Memorandum for DTRA, 24 April 2002.

1-9. Platt, N., Warner, S., and Heagy, J. F., 2002: *User-Oriented Measure of Effectiveness Values for Model Predictions of the Transport and Dispersion of Pollutants Inside a Building*, IDA Memorandum for DTRA, 19 June 2002 and Platt N., Warner, S., and Heagy, J. F., Application of two-dimensional user-oriented measure of effectiveness to interior building releases. *Sixth Annual George Mason University Transport and Dispersion Modeling Workshop,* Fairfax, VA, 19 pp, July 2002

1-10. Warner, S., Platt N., Heagy, J. F., Bradley, S., Bieberbach, G., Sugiyama, G., Nasstrom, J. S., Foster, K. T., and Larson, D., 2001: Model intercomparison with user-oriented measures of effectiveness. *Fifth Annual George Mason University Transport and Dispersion Modeling Workshop,* Fairfax, VA, 15 pp, July 2001.

1-11. Warner, S., Heagy, J. F., Platt, N., Larson, D., Sugiyama, G., Nasstrom, J. S., Foster, K. T., Bradley, S., and Bieberbach, G., 2001: *Evaluation of Transport and Dispersion Models: A Controlled Comparison of Hazard Prediction and Assessment Capability (HPAC) and National Atmospheric Release Advisory Center (NARAC) Predictions*, IDA Paper P-3555, 251 pp, May 2001. (Available electronically [DTIC STINET ada391555] or on CD via e-mail request to Steve Warner at swarner@ida.org or a mail request to Steve Warner, Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, Virginia 22311-1882.)

1-12. Warner, S., Platt, N., and Heagy, 2001: *Application of User-Oriented MOE to HPAC Probabilistic Predictions of* Prairie Grass *Field*, IDA Paper P-3586, 300 pp, May 2001. (Available electronically [DTIC STINET ada391653] or on CD via e-mail request to Steve Warner at swarner@ida.org or a mail request to Steve Warner, Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, Virginia 22311-1882.)

1-13. Warner, S. and Platt, N., 2003: *Analyses in Support of Initial Validation of Urban HPAC: Comparisons to* Urban 2000 *Observations,* IDA Document D-2870N, 119 pp, August 2003. (Available via e-mail request to Steve Warner at swarner@ida.org or a mail request to Steve Warner, Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, Virginia 22311-1882.)

1-14. Warner, S., Platt, N., and Heagy, J. F., 2004: Comparison of transport and dispersion model predictions of the *Urban 2000* field experiment. *J. Appl. Meteor*., **43**: 829-846.

1-15. Heagy, J. F., Platt, N., and Warner, S., 2003: "Analysis and Measures of Effectiveness Values for Predictions of the DF-5," in *Final Report on the DISCRETE FURY Test Program*, for DTRA by Applied Research Associates, Inc. (Report # ARA-LR-3.03-001), Shock Physics Division, P.O. Box 5388, Albuquerque, NM 87185, August 2003.

1-16. Mosca, S., Graziani, G., Klug, W., Bellasio, R., and Bianconi, R., 1998: A Statistical Methodology for the Evaluation of Long-Range Dispersion Models: An Application to the *ETEX* Exercise. *Atmos. Environ.*, 32 (24), 4307-4324.

CHAPTER 2

RESULTS AND DISCUSSION: AREA-BASED MOE VALUES AND COMPARISONS

2. RESULTS AND DISCUSSION: AREA-BASED MOE VALUES AND COMPARISONS

This chapter describes our calculations of area-based MOE values for the 46 sets of predictions of *ETEX*. In order to convert nominal MOE values into area-based MOE values, some way of assessing both the observed and predicted areas must be used. Two different procedures for creating area-based MOE values – Voronoi "weighting" and interpolation via Delaunay triangulation – are discussed in this chapter. Finally, comparisons of the area-based MOE values and nominal MOE values are described.[1] These comparisons focus on rankings based on the previously defined OSF and RWFMS scoring functions.

A. AREA-BASED MOE VALUES FOR PREDICTIONS OF *ETEX*: VORONOI WEIGHTING

This section describes a mathematical procedure that assigns weights to each sampler location. These weights correspond to an assessment of the area represented by each sampler. Samplers placed closer together account for less area than samplers placed further apart.

1. Voronoi Weighting

Given a set of sampler locations across Europe, a Voronoi diagram can be created by dividing Europe into regions such that each region corresponds to one of the sampler locations and all of the points in a given region are closer to the corresponding sampler location than to any other sampler location. Consider the SF_6 sampler locations shown in Figure 2-1. Figure 2-1a shows this region of Europe with latitude and longitude coordinates and Figure 2-1b transforms this picture into Cartesian coordinates with "pseudo-Universal Transverse Mercator" (UTM) kilometers east-west and kilometers north-south appropriately, considering the Earth's curvature. The details of this transformation are described in Appendix B.

---

1 Nominal MOE value computations were previously described [Ref. 2-1].

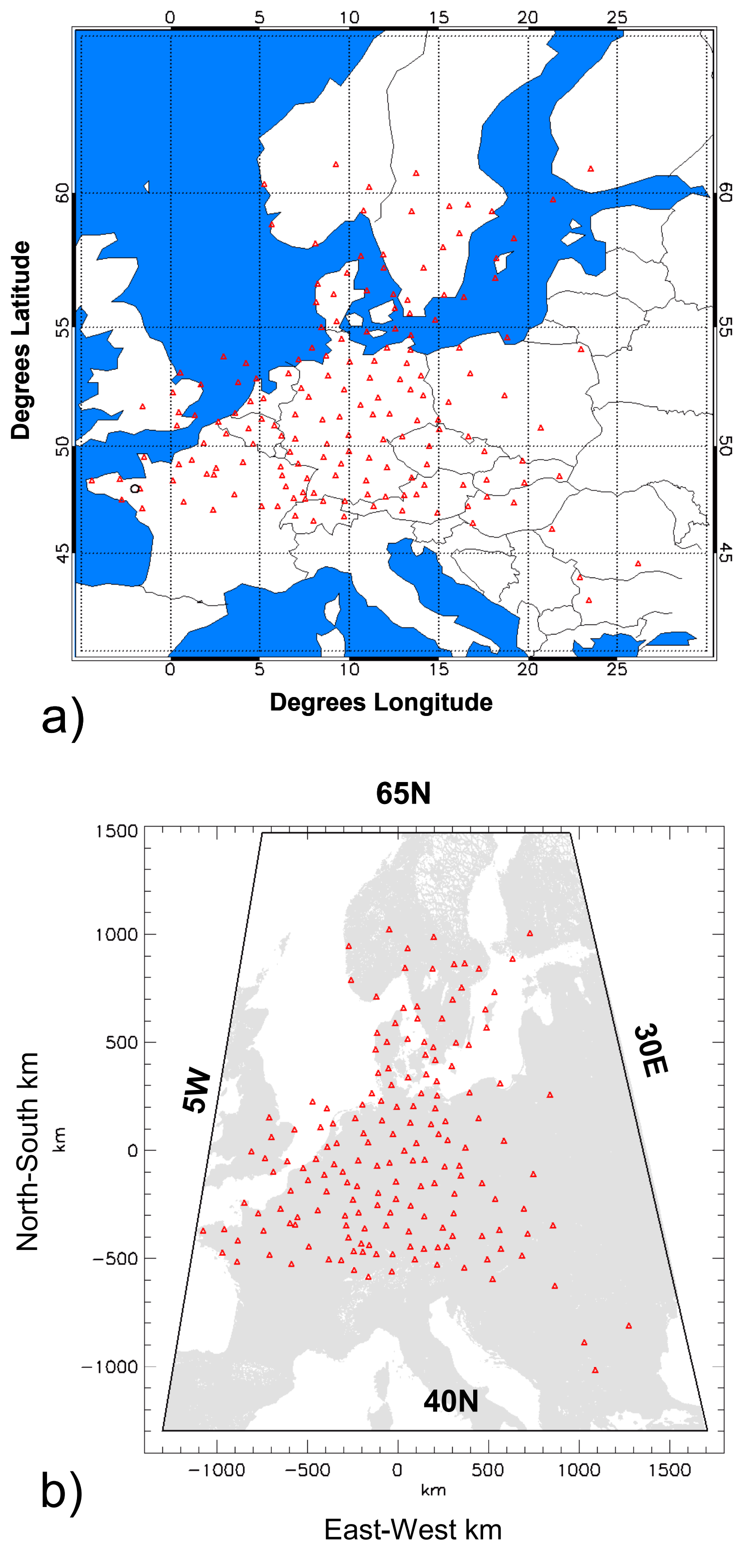


Figure 2-1. ***ETEX*** **Sampler Locations Across Europe: Red Open Triangles Correspond to Sampler Locations and the Black Open Circle Corresponds to the Release Location: a) x-axis is in Degrees Longitude and y-axis is in Degrees Latitude b) x-axis is in East-West kilometers and y-axis is in North-South kilometers.**

Figure 2-2 shows the associated Voronoi diagram for the *ETEX* sampler locations. This Voronoi diagram was generated using algorithms resident in Interactive Data Language (IDL) software.[2] Note that many of the polygons are closed, but some extend past the European domain. These polygons that extend beyond Europe could result in very large areas – areas in which for the most part there are no samplers and generally little or no predicted or observed SF_6. When assessing area-based MOE values for the lowest thresholds (or contours), we used the following two procedures to truncate the size of these potentially large regions that had little sampler coverage. First, the shaded region in Figure 2-2 (and Figure 2-1b) was used as the complete domain over which the calculations would be done. That is, in this scheme, those polygons that extend beyond this domain are truncated – i.e., "natural geographic extent clipping." Then, the Voronoi diagram can be used to assign area sizes to each polygon or truncated polygon. These area sizes (a_i) are then used to directly weight each prediction/observation comparison and in this way an approximate area-based MOE value can be created that does not rely on area interpolation. We refer to MOE values computed in this way as "Voronoi weighted with clipping" (VW_c). Alternatively, one can truncate all Voronoi polygons at the size of the nth percentile polygon size. Figure 2-3 illustrates the 80th percentile area size for this *ETEX* Voronoi diagram. Then, using the polygon area sizes (a_1, a_2, a_3,… a_i), and the truncated 80th percentile polygon area size (a_{80th}) for the largest (20 percent of the) polygons, one can again create an approximate area-based MOE value does not rely on area interpolation. We refer to MOE values computed in this way as "Voronoi weighted with 80th percentile clipping" (VW_{80}).

2. Area-Based MOE Values Using Voronoi Weighting

The MOE (as mentioned in Chapter 1) is described below in terms of false negative (A_{FN}), false positive (A_{FP}), observed (A_{OB}), and predicted (A_{PR}) regions.

$$MOE = (x, y) = \left(\frac{A_{OV}}{A_{OB}}, \frac{A_{OV}}{A_{PR}} \right) = \left(\frac{A_{OB} - A_{FN}}{A_{OB}}, \frac{A_{PR} - A_{FP}}{A_{PR}} \right) = \left(1 - \frac{A_{FN}}{A_{OB}}, 1 - \frac{A_{FP}}{A_{PR}} \right). \qquad (2\text{-}1)$$

An additional algebraic step, substituting $A_{FN} + A_{OV} = A_{OB}$ and $A_{FP} + A_{OV} = A_{PR}$, leads to an MOE description in terms of A_{FN}, A_{FP}, and A_{OV}.

$$MOE = \left(1 - \frac{A_{FN}}{A_{OB}}, 1 - \frac{A_{FP}}{A_{PR}} \right) = \left(1 - \frac{A_{FN}}{A_{FN} + A_{OV}}, 1 - \frac{A_{FP}}{A_{FP} + A_{OV}} \right). \qquad (2\text{-}2)$$

2 See Reference 2-2.

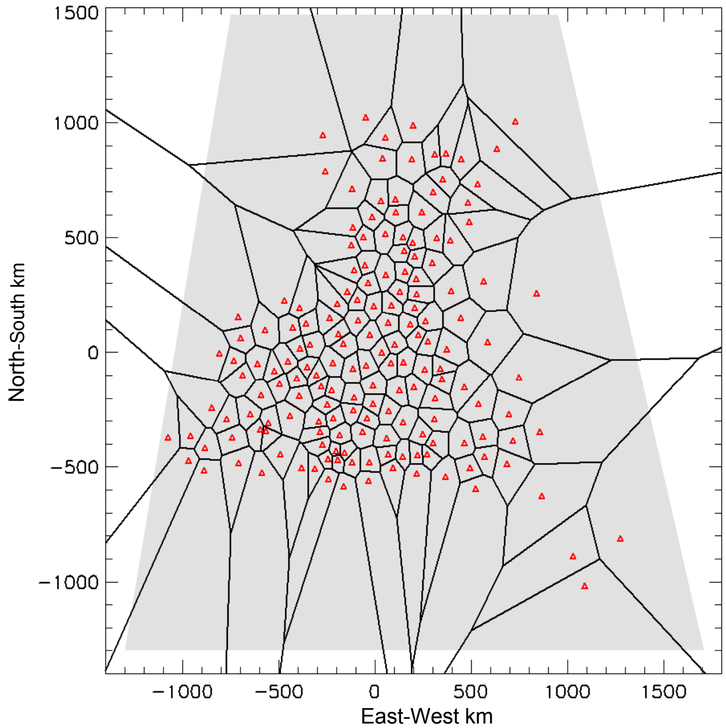


Figure 2-2. Voronoi Diagram Based On *ETEX* Sampler Locations: Red Open Triangles Correspond to Sampler Locations, Gray Shading Corresponds to Computational Domain.

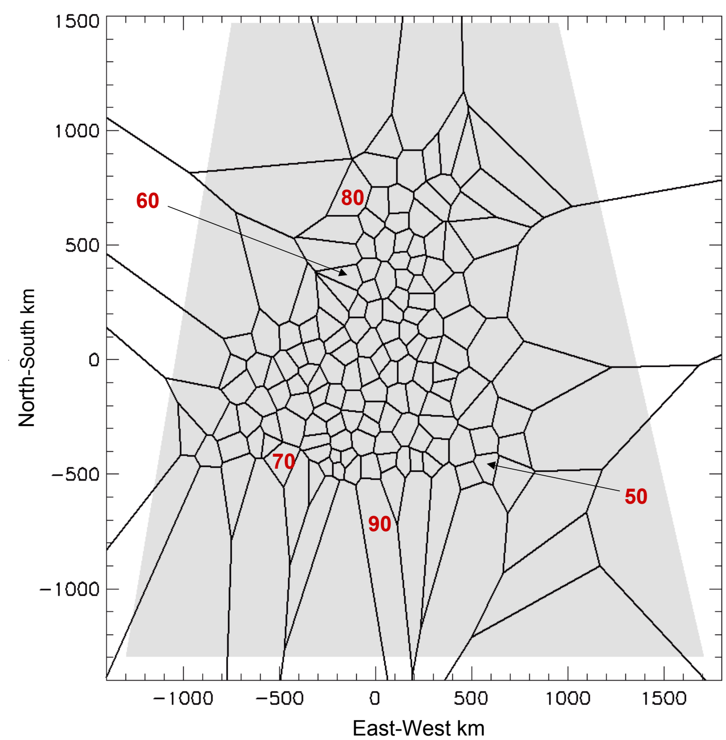


Figure 2-3. Voronoi Diagram Based On *ETEX* Sampler Locations: Red Labels Refer to the Polygon Area Size as a Percentile of All Polygon Area Sizes (e.g., "80" implies 80th percentile).

Although the MOE has been described above in terms of the three areas (or regions) A_{FN}, A_{FP}, and A_{OV}, it is not necessary to have actual physical areas to compute the components of the MOE. Rather, A_{FN}, A_{FP}, and A_{OV} can be computed directly from the predictions and field trial observations paired in space and time. For example, for a concentration-based MOE, the false positive region is the concentration predicted in a region but not observed. Therefore, for A_{FP}, one first considers all of the samplers at which the prediction is of greater value than the observation. Next, one sums the differences between the predicted and observed concentrations at those samplers. Based on the samplers that contained observed values that were larger than the predicted values, one can similarly compute A_{FN}. A_{OV} is calculated by considering all samplers and summing the concentrations associated with the minimum predicted or observed value. Restating the above mathematically, let

$$\begin{aligned} A_{OV}(i) &= \min\{predicted(i), observed(i)\} \\ A_{FN}(i) &= \max\{observed(i) - A_{OV}(i), 0\} \quad , \\ A_{FP}(i) &= \max\{predicted(i) - A_{OV}(i), 0\} \end{aligned} \tag{2-3}$$

then for N observation/prediction pairings

$$\begin{aligned} A_{OV} &= \sum_{i=1}^{N} A_{OV}(i) \\ A_{FN} &= \sum_{i=1}^{N} A_{FN}(i) \, . \\ A_{FP} &= \sum_{i=1}^{N} A_{FP}(i) \end{aligned} \tag{2-4}$$

These estimates can be made on a linear scale or on a logarithmic scale. If dosage information were used in place of concentrations, an analogous procedure could be used to compute dosage-based MOE values.

In addition to the more general technique described above, one can compute an MOE value based on an identified threshold (e.g., concentration or dosage) of interest as notionally illustrated in Figure 1-2. First, one considers the predictions and observations at each of the samplers. If both the prediction and observation are above the threshold, it is considered overlap at that sampler. If the prediction is below the threshold and the observation is above, a false negative is assessed at that sampler. Similarly, a false positive is assessed when the prediction is above the threshold and the observation is not.

Restating mathematically, given a set of samplers with observations and predictions and a threshold, *T*, one can partition this set into four subsets – *OV*, *FN*, *FP*, and *BELOW*:

$$
\begin{aligned}
OV &= \{i \mid observed(i) \geq T \text{ and } predicted(i) \geq T\} \\
FN &= \{i \mid observed(i) \geq T \text{ and } predicted(i) < T\} \\
FP &= \{i \mid observed(i) < T \text{ and } predicted(i) \geq T\} \\
BELOW &= \{i \mid observed(i) < T \text{ and } predicted(i) < T\}
\end{aligned} \quad . \qquad (2\text{-}5)
$$

Then

$$
\begin{aligned}
A_{OV} &= \text{number of elements (samplers) in } OV \\
A_{FN} &= \text{number of elements (samplers) in } FN \\
A_{FP} &= \text{number of elements (samplers) in } FP
\end{aligned} \quad . \qquad (2\text{-}6)
$$

It is possible to modify the above definition of A_{OV} to include the number of elements (samplers) in the *BELOW* set. To be consistent with the conceptual view illustrated in Figure 1-2, A_{OV} was defined as in Eq. 2-6.

For the case of this analysis, observed and predicted 3-hour average concentrations were compared for the 90 hours of SF_6 monitoring during *ETEX*. There are 30 (3 hour) time periods for comparisons and Eq. 2-3 is modified to account for this as shown in Eq. 2-7. Therefore, for polygon *i* and time period, *t*,

$$
\begin{aligned}
A_{OV}(i,t) &= \min\{predicted(i,t), observed(i,t)\} \\
A_{FN}(i,t) &= \max\{observed(i,t) - A_{OV}(i,t), 0\} \\
A_{FP}(i,t) &= \max\{predicted(i,t) - A_{OV}(i,t), 0\}
\end{aligned} \quad . \qquad (2\text{-}7)
$$

To compute Voronoi-weighted values of A_{OV}, A_{FN}, and A_{FP} – and ultimately a Voronoi-weighted MOE value – the Voronoi polygon area sizes (a_i) are included as follows when considering MOE values based on summed concentrations (i.e., marginal differences).

$$
\begin{aligned}
A_{OV}(t) &= \sum_{i=1}^{N} \left(a_i \times A_{OV}(i,t)\right) \\
A_{FN}(t) &= \sum_{i=1}^{N} \left(a_i \times A_{FN}(i,t)\right) \\
A_{FP}(t) &= \sum_{i=1}^{N} \left(a_i \times A_{FP}(i,t)\right)
\end{aligned} \quad . \qquad (2\text{-}8)
$$

Then, after summing for the 30 time periods we have

$$
\begin{aligned}
A_{OV} &= \sum_{t=1}^{30} A_{OV}(t) \\
A_{FN} &= \sum_{t=1}^{30} A_{FN}(t) \, . \\
A_{FP} &= \sum_{t=1}^{30} A_{FP}(t)
\end{aligned} \tag{2-9}
$$

These values can then be used to compute a Voronoi-weighted (area-based) MOE for summed concentrations (also referred to as marginal concentration differences).

A similar procedure is followed to compute Voronoi-weighted values when considering a critical threshold, T. First, we define the delta functions $\delta_{OV}(i,t)$, $\delta_{FN}(i,t)$, $\delta_{FP}(i,t)$ such that

$$
\begin{aligned}
\delta_{OV}(i,t) &= \begin{Bmatrix} 1 & \textit{if observed}(i,t) \geq T \text{ and } \textit{predicted}(i,t) \geq T \\ 0 & \textit{otherwise} \end{Bmatrix} \\
\delta_{FN}(i,t) &= \begin{Bmatrix} 1 & \textit{if observed}(i,t) \geq T \text{ and } \textit{predicted}(i,t) < T \\ 0 & \textit{otherwise} \end{Bmatrix} . \\
\delta_{FP}(i,t) &= \begin{Bmatrix} 1 & \textit{if observed}(i,t) < T \text{ and } \textit{predicted}(i,t) \geq T \\ 0 & \textit{otherwise} \end{Bmatrix}
\end{aligned} \tag{2-10}
$$

The Voronoi-weights are applied as in Eq. 2-11.

$$
\begin{aligned}
A_{OV}(t) &= \sum_{i=1}^{N} \left(a_i \times \delta_{OV}(i,t) \right) \\
A_{FN}(t) &= \sum_{i=1}^{N} \left(a_i \times \delta_{FN}(i,t) \right) . \\
A_{FP}(t) &= \sum_{i=1}^{N} \left(a_i \times \delta_{FP}(i,t) \right)
\end{aligned} \tag{2-11}
$$

Finally, the 30 time periods are summed as in Eq. 2-9.

Figure 2-4 presents the MOE values associated with predictions of 3-hour average concentrations[3] and based on a threshold of 0.01 ng m^{-3}. The MOE values of Figure 2-4 provide information on model performance with respect to predicting the *area size* of 3-hour average concentrations above 0.01 ng m^{-3}. The numbers in Figure 2-4 correspond to the model number (Table 1-1) with the blue labels referring to the 100 series (e.g., the

---

[3] The sample collection time was 3 hours and thus represents the highest time resolution associated with these data.

blue “12” implies model 112) and the red labels referring to the 200 series (e.g., the red “8” implies model 208). Figures 2-4a and 2-4b present results based on the VW_{80} and VW_c procedures, respectively.

Most of the 46 model predictions cluster in an elliptical region that straddles the diagonal (the dashed light purple line). Exceptions to this are models 117, 129, 130, 212, and 214. There are differences between the MOE values computed by the two VW procedures, but the overall character of the plot is similar. Recall that an MOE value on this diagonal implies equal sizes of the observed and predicted region – in this case, area size. The variation in MOE performance for the different models appears roughly perpendicular to the diagonal line. The implication of this variation is simply that some models led to over-predictions (those below the diagonal) and some led to under-predictions (those above the diagonal).

Figures 2-5 (a and b) and 2-6 (a and b) present area based VW MOE values for the 46 model predictions of *ETEX* at threshold values of 0.1 and 0.5 ng m^{-3}, respectively. Figure 2-6 shows that at the highest examined threshold (0.5 ng m^{-3}), most model predictions tend to over-predict (lie below the diagonal) the number of locations that exceed the threshold and relative MOE performance is worse than at the lower thresholds. Table 2-1 identifies the top ranked model predictions as judged by the OSF as well as the rankings (out of 46) of SCIPUFF and ARAC. Rankings are identified for the area-based (VW) MOE values at each of three thresholds – i.e., contour levels. Values in red correspond to VW_{80}-based results, values in blue correspond to VW_c-based results, and values in black imply identical results were found for both VW-based procedures. No single model dominated the top ranking. Rankings based on RWFMS with C_{FN} = C_{FP} = 1.0 – RWFMS(1,1) – are shown in Table 2-2. Rankings based on RWFMS with C_{FN} = 5.0 and C_{FP} = 0.5 – RWFMS(5,0.5) – are shown in Table 2-3. Complete rankings can be found later in this chapter.

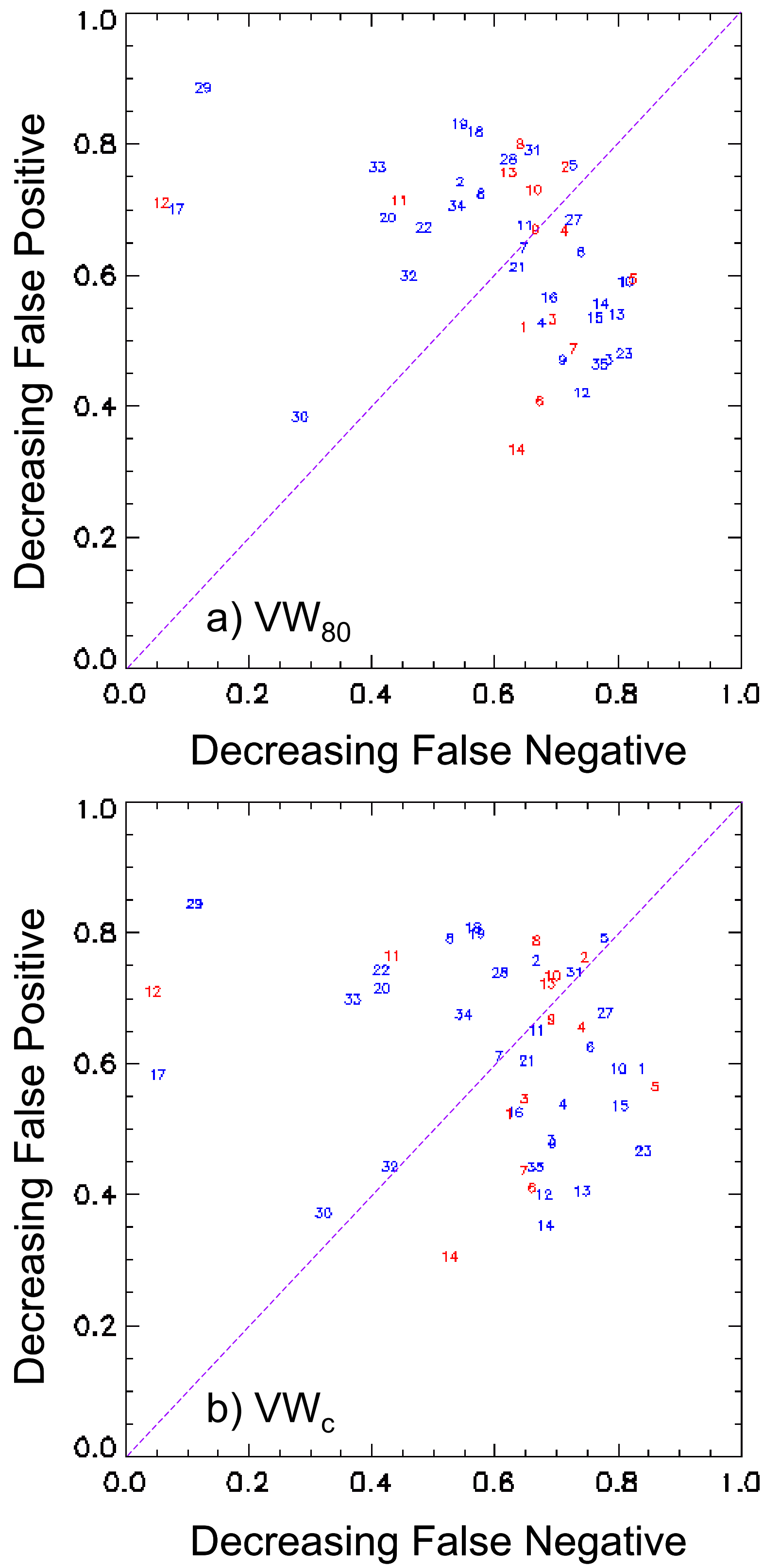


Figure 2-4. Area-Based (Voronoi Weighted) MOE Values for a 0.01 ng m^{-3} Threshold: a) VW_{80} and b) VW_c

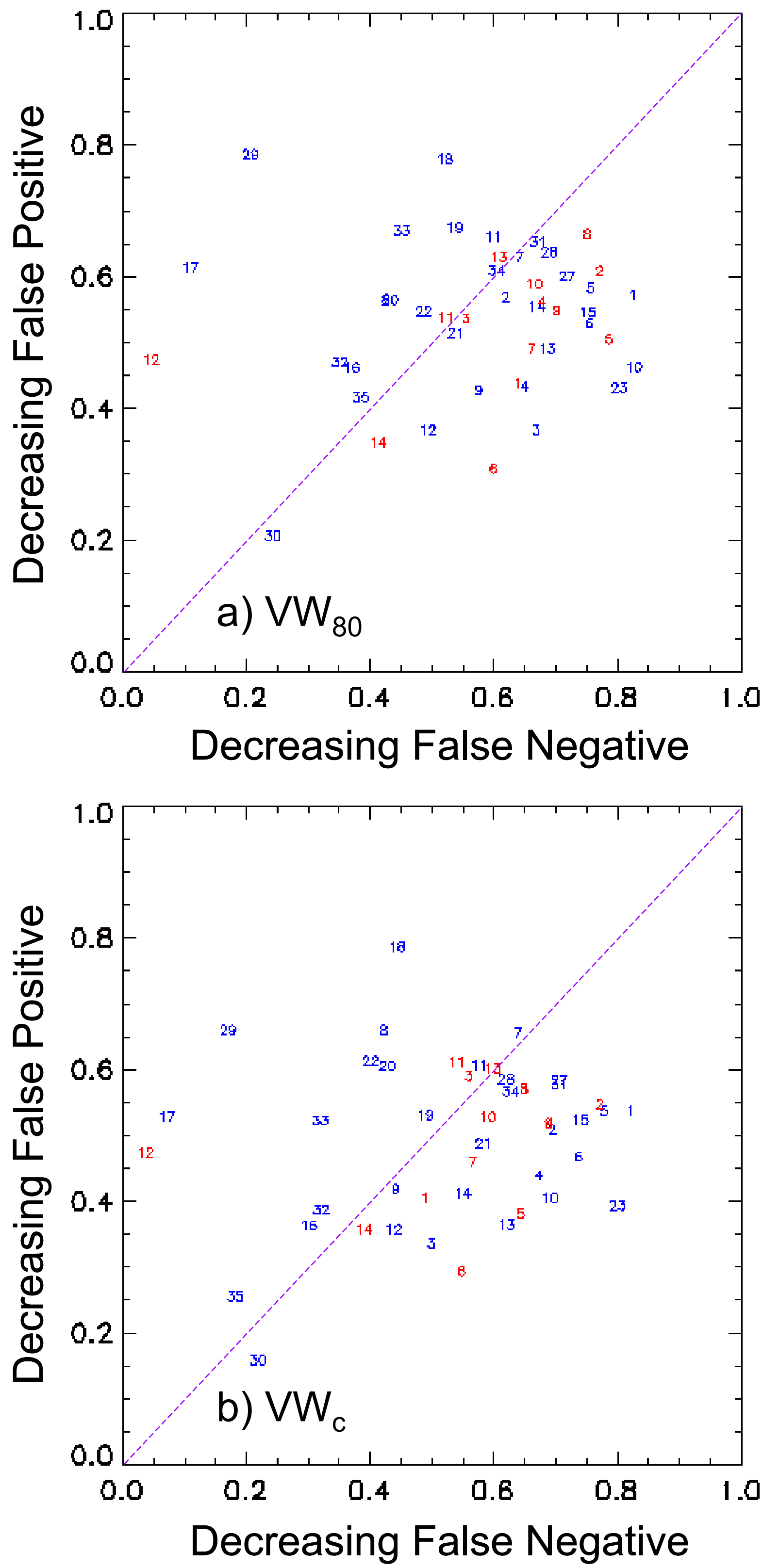
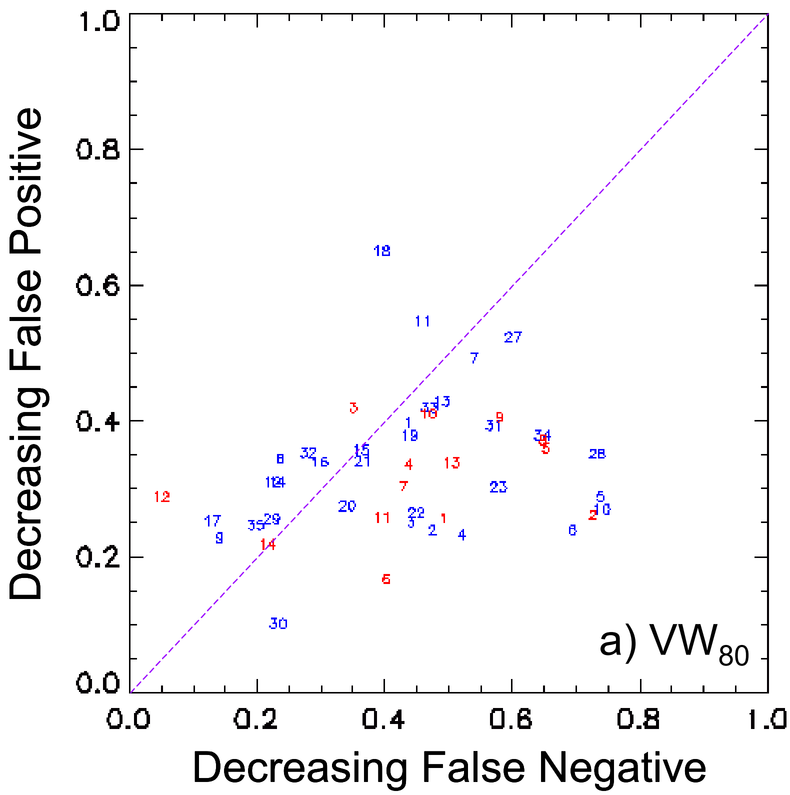


Figure 2-5. Area-Based (Voronoi Weighted) MOE Values for a 0.1 ng m^{-3} Threshold: a) VW_{80} and b) VW_c



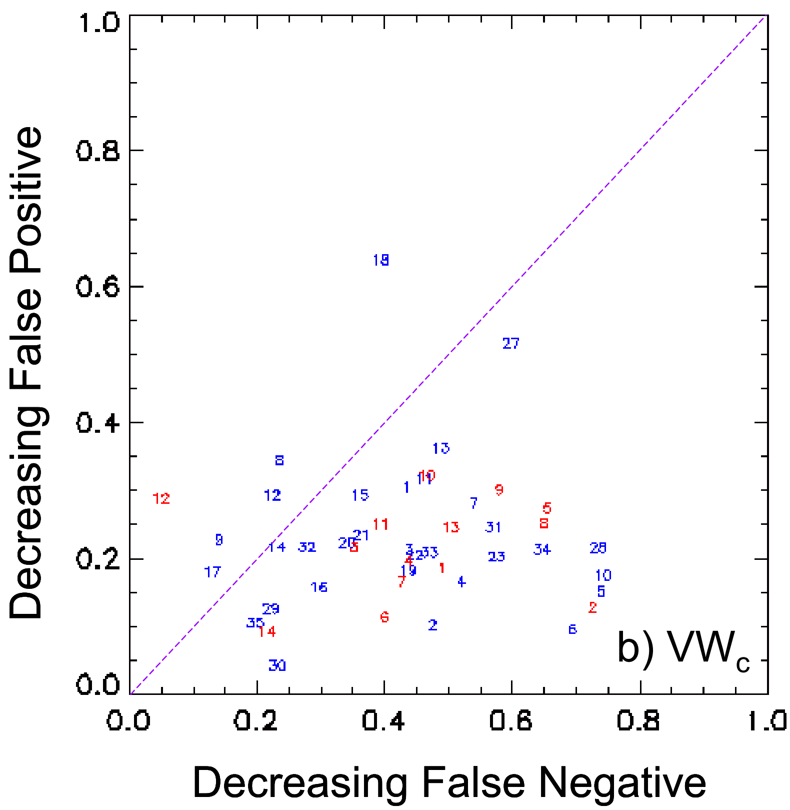


Figure 2-6. Area-Based (Voronoi Weighted) MOE Values for a 0.5 ng m^{-3} Threshold: a) VW_{80} and b) VW_c

Table 2-1. Top-Ranked Model and Rankings of SCIPUFF and ARAC Based on OSF – VW_{80} and VW_c

Rank	0.01 ng m^{-3}	0.1 ng m^{-3}	0.5 ng m^{-3}
1	Canadian Meteorological Centre - CMC (105)	Swedish Meteorological and Hydrological Office - SMHI (208) / German Weather Service – DWD (107)	Lawrence Livermore National Laboratory – ARAC (127)
Model	**0.01 ng m^{-3}**	**0.1 ng m^{-3}**	**0.5 ng m^{-3}**
SCIPUFF	24 / 21	30 / 23	28 / 29
ARAC	5 / 4	7 / 3	1

Table 2-2. Top-Ranked Model and Rankings of SCIPUFF and ARAC Based on RWFMS(1,1) – VW_{80} and VW_c

Rank	0.01 ng m^{-3}	0.1 ng m^{-3}	0.5 ng m^{-3}
1	Canadian Meteorological Centre - CMC (105)	Swedish Meteorological and Hydrological Office - SMHI (208) / German Weather Service – DWD (107)	Lawrence Livermore National Laboratory – ARAC (127)
Model	**0.01 ng m^{-3}**	**0.1 ng m^{-3}**	**0.5 ng m^{-3}**
SCIPUFF	24 / 21	29 / 21	25 / 20
ARAC	5 / 4	7 / 3	1

Table 2-3. Top-Ranked Model and Rankings of SCIPUFF and ARAC Based on RWFMS(5,0.5) – VW_{80} and VW_c

Rank	0.01 ng m^{-3}	0.1 ng m^{-3}	0.5 ng m^{-3}
1	Danish Meteorological Institute - DMI (205)	Institute of Meteorology and Physics, University of Wien - IMP (101)	Swedish Meteorological and Hydrological Office - SMHI (128)
Model	**0.01 ng m^{-3}**	**0.1 ng m^{-3}**	**0.5 ng m^{-3}**
SCIPUFF	27 / 22	32 / 23	30 / 31
ARAC	12 / 6	10 / 7	8 / 2

B. AREA-BASED MOE VALUES FOR PREDICTIONS OF *ETEX*: INTERPOLATION VIA DELAUNAY TRIANGULATION

Given values at some discrete (perhaps irregular) set of samplers, the process of interpolation provides intermediate values on some regular grid of points. The resulting regular grid of functional values could be used to obtain contours of "hazard" areas (areas within a critical threshold contour) or calculate MOE values based on interpolated areas. Interpolated values can also be used to display "hazard" areas according to dosage intensity as is shown in Figure 2-7.

Interpolation procedures can be carried out either in linear or logarithmic space. When interpolating actual plume concentrations or dosages varying over orders of magnitude, one might favor interpolation schemes in logarithmic space.

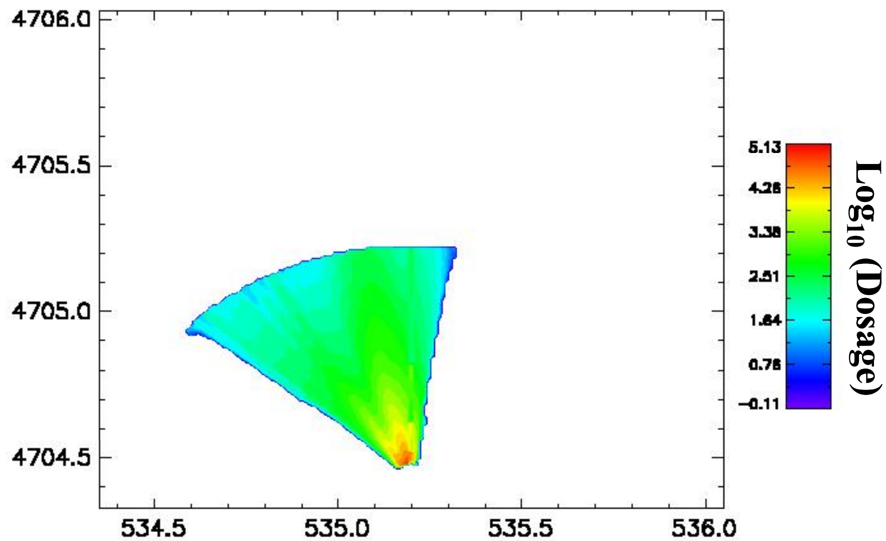


Figure 2-7. Example Dosages (in mg-sec/m^3) for *Prairie Grass* Field Trial 43 Using Delaunay Triangulation Procedure to Perform Interpolation to a Regular Grid of Points [from Ref. 2-3]

1. Delaunay Triangulation

The Delaunay triangulation procedure is useful for the interpolation, analysis, and visual display of irregularly, discretely gridded data. From a set of discrete points (sampler coordinates), a planar triangulation is formed, satisfying the property that the circumscribed circle of any triangle in the triangulation contains no other vertices in its interior.[4] For any point that is within some triangle (formed via Delaunay triangulation),

[4] The Delaunay triangulation procedure is closely related to the procedure followed to create Voronoi polygons [Ref. 2-4].

a linear interpolation routine using values at the vertices of the triangle is used to calculate the value at that point. Delaunay triangulation is efficiently implemented in IDL and forms a core interpolation routine for display of irregularly gridded data.

We used the above procedure in the following way. First, we transformed the data (observations and predictions) logarithmically and then followed the above procedure. This routine was applied with a resolution of 2 km by 2 km corresponding to 1001×1001 grid points. The displays reported in Figure 2-8 [Ref. 2-1] are based on the logarithmic transformation of the data followed by Delaunay triangulation and linear interpolation as described above. The adopted procedure, while simple and yielding some perhaps less visually pleasing sharp edges, appeared to be robust and necessarily maintains the actual observed values at the sampler locations – this would not be not true for many fitting procedures.[5]

2. Area-Based MOE Values Using Interpolation Via Delaunay Triangulation (IDT)

Figure 2-9 presents the MOE values associated with predictions of 3-hour average concentrations[6] and based on a threshold of 0.01 ng m^{-3}. The MOE values of Figure 2-9 provide information on model performance with respect to predicting the *area size* of 3-hour average concentrations above 0.01 ng m^{-3}. The numbers in Figure 2-9 correspond to the model number (Table 1-1) with the blue labels referring to the 100 series (e.g., the blue "12" implies model 112) and the red labels referring to the 200 series (e.g., the red "8" implies model 208).

An ellipse has been drawn in Figure 2-9 to highlight the result that most of the 46 models led to MOE values in a relatively similar location in the MOE space. Only eight of the model predictions lie outside of this (arbitrary) ellipse (117, 120, 129, 130, 132, 206, 212, and 214). For the 39 model predictions that led to MOE values within the ellipse, it can be seen that they straddle the diagonal line. As described for the VW area-based MOE values, the variation in IDT area-based MOE performance for the different models within the ellipse appears roughly perpendicular to the diagonal line. The implication of this variation is simply that some models led to over-predictions (those below the diagonal) and some led to under-predictions (those above the diagonal).

5 A few other interpolation techniques were briefly examined in Ref. 2-1.

6 The sample collection time was three hours and thus represents the highest time resolution associated with these data.

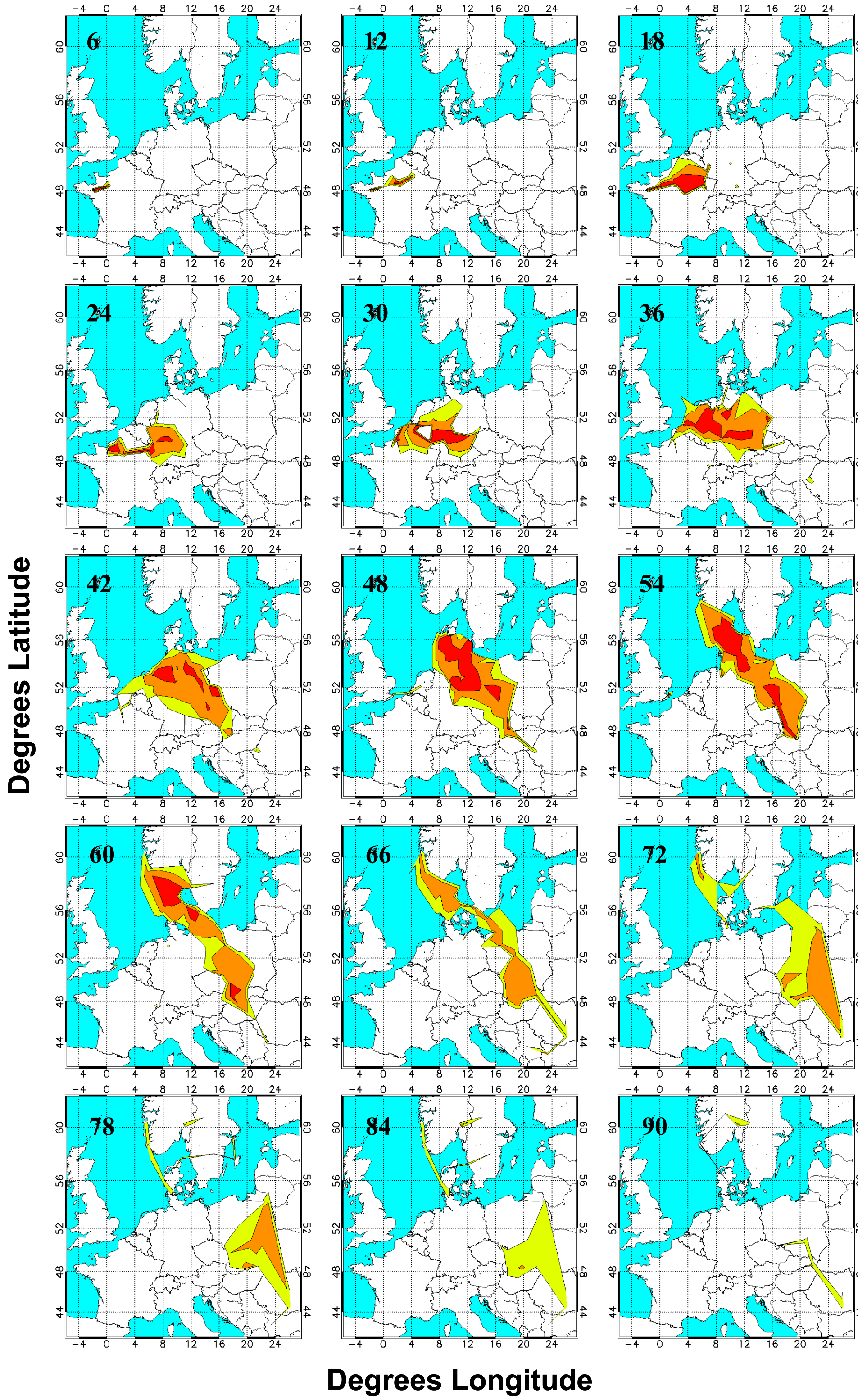


Figure 2-8. Observed PMCH Concentrations Across Europe. Plots Display Contours from 6 Hours After the Release for the Upper Left Plot to 90 Hours After the Release for the Lower Right Plot in Increments of 6 Hours. Contours are 0.01, 0.1, and 0.5 ng m^{-3}. Bold numbers on individual plots correspond to the last hour of the given 6-hour period.

A few models, like 105, 107, 111, and 121 resulted in MOE values very near the diagonal, implying little bias in the prediction of the area size that exceeds the 0.01 ng m^{-3} threshold (neither an over- nor under-prediction on average).[7]

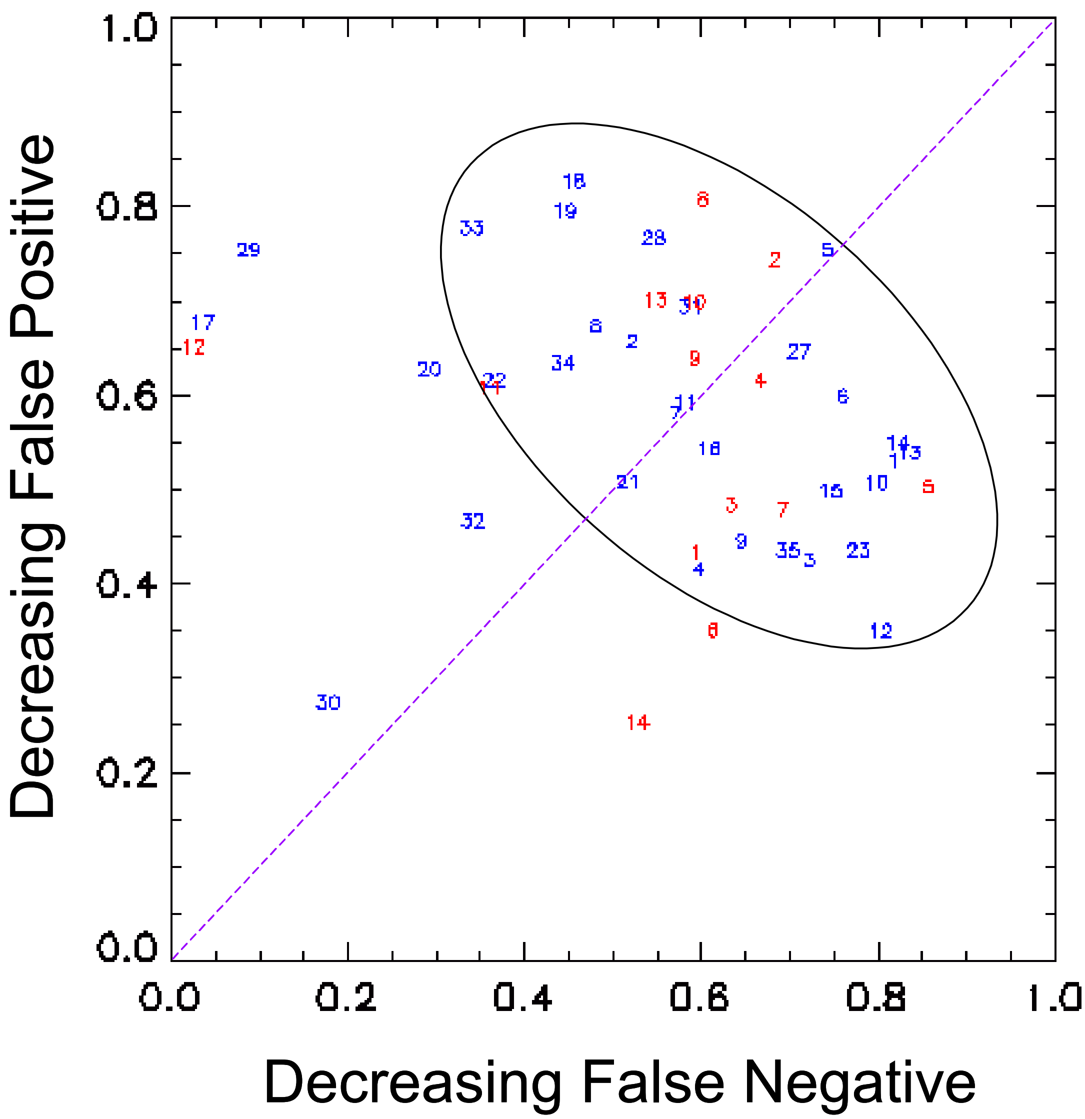


Figure 2-9. 3-Hour Average Concentration Area-Based (IDT) Threshold (0.01 ng m^{-3}) MOE Values for 46 ATMES II Participants. Blue Numbered Labels Refer to Series 100 Models (e.g., "19" implies model 119) and Red Numbered Labels Refer to Series 200 Models (e.g., "13" implies model 213).

---

7 The results described here are similar to those described for the *nominal* 0.01 ng m^{-3} threshold based MOE [Ref. 2-1].

Figure 2-10 and 2-11 present area-based MOE values (interpolation via Delaunay triangulation - IDT) for thresholds of 0.1 and 0.5 ng m^{-3}, respectively. It is apparent (Figure 2-11) that at the 0.5 ng m^{-3} level, most of the models over-predicted the "hazard" area.

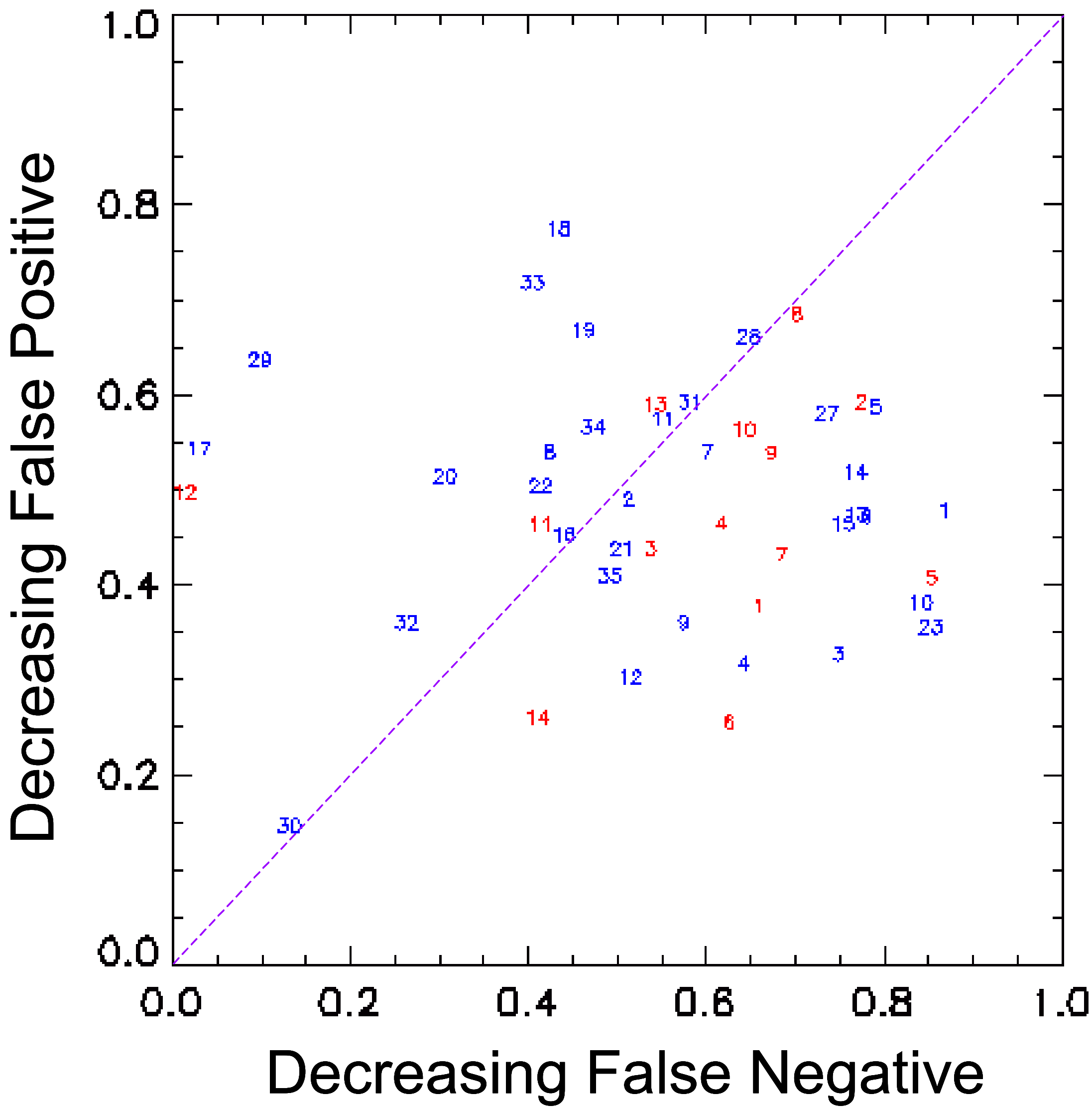


Figure 2-10. 3-Hour Average Concentration Area-Based (IDT) Threshold (0.1 ng m^{-3}) MOE Values for 46 ATMES II Participants. Blue Numbered Labels Refer to Series 100 Models (e.g., "19" implies model 119) and Red Numbered Labels Refer to Series 200 Models (e.g., "13" implies model 213).

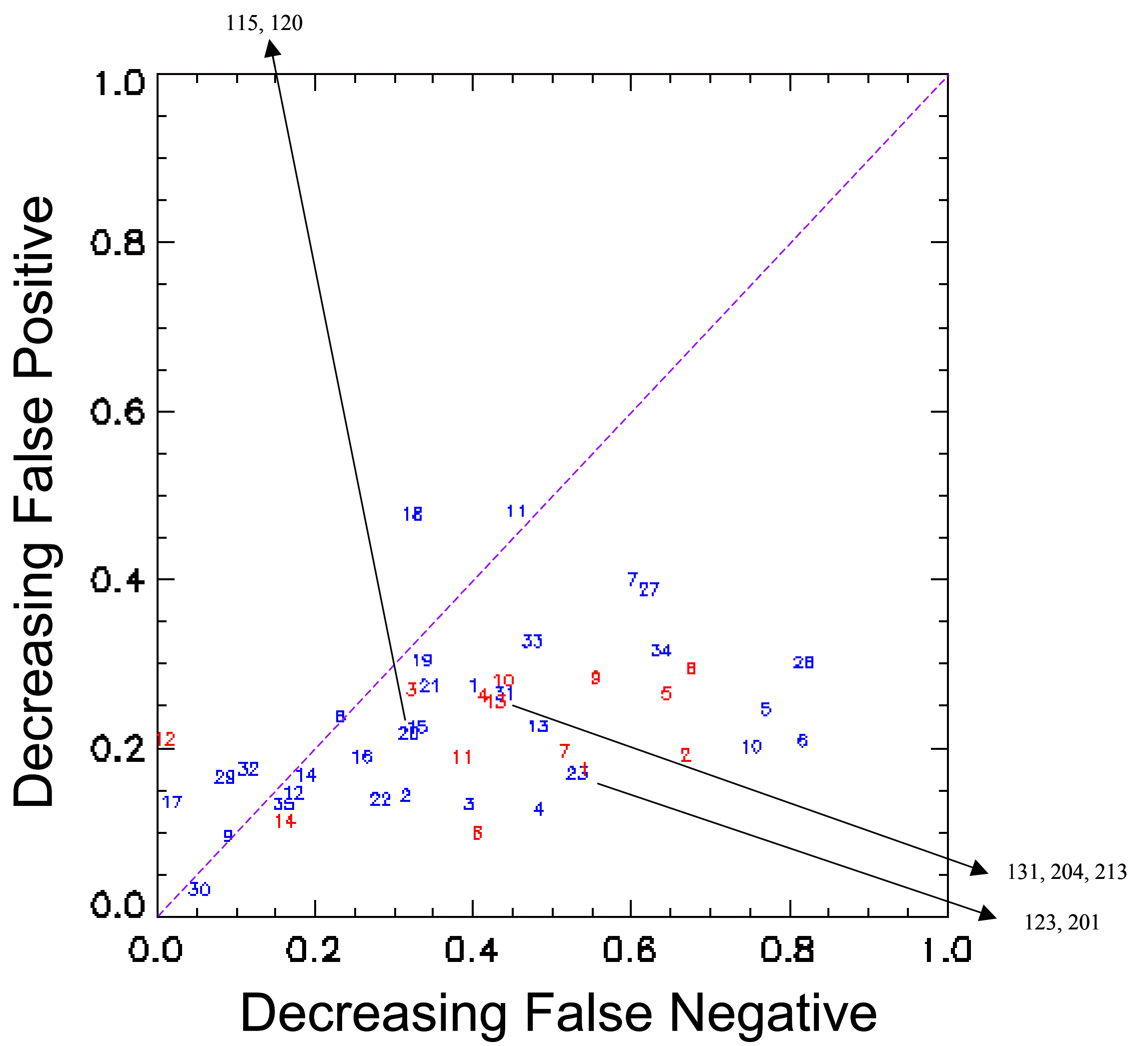


Figure 2-11. 3-Hour Average Concentration Area-Based (IDT) Threshold (0.5 ng m^{-3}) MOE Values for 46 ATMES II Participants. Blue Numbered Labels Refer to Series 100 Models (e.g., "19" implies model 119) and Red Numbered Labels Refer to Series 200 Models (e.g., "13" implies model 213).

Table 2-4 identifies the top ranked model predictions as judged by the OSF as well as the rankings (out of 46) of SCIPUFF and ARAC. Rankings are identified for the area-based (IDT) MOE values at each of three thresholds – i.e., contour levels. No single model dominated the top ranking.[8] Rankings based on RWFMS with $C_{FN} = C_{FP} = 1.0$ – RWFMS(1,1) – are shown in Table 2-5. Rankings based on RWFMS with $C_{FN} = 5.0$ and $C_{FP} = 0.5$ – RWFMS(5,0.5) – are shown in Table 2-6. The next section of this chapter compares these rankings with nominal and VW-based MOE rankings.

8 Complete rankings can be found later in this chapter.

Table 2-4. Top-Ranked Model and Rankings of SCIPUFF and ARAC Based on OSF – IDT

Rank	0.01 ng m^{-3}	0.1 ng m^{-3}	0.5 ng m^{-3}
1	Canadian Meteorological Centre - CMC (105)	Swedish Meteorological and Hydrological Office - SMHI (208)	Lawrence Livermore National Laboratory – ARAC (127)
Model	**0.01 ng m^{-3}**	**0.1 ng m^{-3}**	**0.5 ng m^{-3}**
SCIPUFF	33	31	25
ARAC	4	5	1

Table 2-5. Top-Ranked Model and Rankings of SCIPUFF and ARAC Based on RWFMS(1,1) – IDT

Rank	0.01 ng m^{-3}	0.1 ng m^{-3}	0.5 ng m^{-3}
1	Canadian Meteorological Centre - CMC (105)	Swedish Meteorological and Hydrological Office - SMHI (208)	German Weather Service – DWD (107)
Model	**0.01 ng m^{-3}**	**0.1 ng m^{-3}**	**0.5 ng m^{-3}**
SCIPUFF	32	30	21
ARAC	4	5	2

Table 2-6. Top-Ranked Model and Rankings of SCIPUFF and ARAC Based on RWFMS(5,0.5) – IDT

Rank	0.01 ng m^{-3}	0.1 ng m^{-3}	0.5 ng m^{-3}
1	Danish Meteorological Institute - DMI (205)	Institute of Meteorology and Physics, University of Wien - IMP (101)	Swedish Meteorological and Hydrological Office - SMHI (128)
Model	**0.01 ng m^{-3}**	**0.1 ng m^{-3}**	**0.5 ng m^{-3}**
SCIPUFF	32	29	25
ARAC	11	10	6

In this section, MOE computations were based on interpolations that were accomplished after logarithmic transformation (as previously described). We also examined MOE values based on interpolation applied to the untransformed or "raw" observations and predictions. Comparisons of the two procedures showed some changes in terms of model rankings but overall only small changes. For instance, the median absolute difference in OSF model rankings when comparing the "logarithmic transformation" to "no transformation" interpolation procedures for the 0.01, 0.1, and 0.5 ng m^{-3} were 3, 3, and 2.5, respectively.

C. COMPARISONS OF MODEL RANKINGS: VORONOI WEIGHTED, INTERPOLATION VIA DELAUNAY TRIANGULATION, AND NOMINAL

Table 2-7 compares nominal and area-based MOE rankings for OSF. The nominal MOE rankings were previously computed in Ref. 2-1. The rankings are relatively robust to the different techniques (nominal and area-based) used to compute the MOE values. For example, of the top 10 OSF-ranked models based on the nominal MOE, 8, 7, and 8 appear in the top ten based on the VW_{80}, VW_c, and IDT MOE values, respectively. Similarly, of the bottom 10 OSF-ranked models based on the nominal MOE, 9, 8, and 9 appear in the bottom 10 based on the VW_{80}, VW_c, and IDT MOE values, respectively.

Figure 2-12 shows a histogram of the changes in rankings that result from subtracting the area-based MOE rankings from the nominal MOE rankings. First, the mean and median absolute difference in rankings varies from 1.8 and 1 for VW_{80} based rankings to 4.3 and 2 for VW_c based rankings (and 3 and 2 for IDT based rankings) as shown in the figure. For perspective, we simulated the random ordering of 46 entities and found the mean and median absolute ranking differences were 15.3 and 13.5. The small changes in rankings shown in Figure 2-12 suggest the relative robustness of model prediction OSF rankings given the various computational techniques – nominal, VW_{80}, VW_c, and IDT. Model 102 – BMRC from Table 1-1 – ranked 33 using the nominal MOE and OSF, moves up 11 places to 22, 26 places to 7, and 13 places to 20 for the VW_{80}, VW_c, and IDT MOE values, respectively. It appears that a large under-prediction that occurred early in the release for model 102 led to this result. At the earlier times and shorter ranges, the corresponding polygon sizes (and hence weighting) and interpolated area sizes are, in general, smaller. Therefore, for model 102, the increased importance of the longer range and later-in-time prediction/observation comparisons that were

Table 2-7. Comparisons of Model Rankings Based on OSF for a 0.01 ng m^{-3} Threshold: Nominal and Area-Based – VW_{80}, VW_c, and IDT – MOE Values

Rank	Model	Nominal	Model	VW_{80}	Model	VW_c	Model	IDT
1	202	0.358	105	0.362	105	0.308	105	0.360
2	105	0.361	202	0.378	202	0.356	202	0.415
3	208	0.388	131	0.401	131	0.379	208	0.449
4	127	0.389	208	0.418	127	0.395	127	0.461
5	128	0.397	127	0.421	208	0.402	106	0.473
6	210	0.413	210	0.436	210	0.409	114	0.491
7	131	0.420	128	0.442	102	0.415	113	0.495
8	101	0.420	204	0.447	213	0.423	101	0.509
9	205	0.420	205	0.450	204	0.439	210	0.511
10	114	0.424	213	0.453	101	0.444	128	0.513
11	106	0.427	106	0.453	106	0.453	131	0.516
12	110	0.431	110	0.457	110	0.460	204	0.517
13	204	0.439	101	0.458	209	0.461	205	0.525
14	118	0.441	118	0.470	205	0.465	110	0.538
15	209	0.445	209	0.478	128	0.474	213	0.545
16	107	0.451	111	0.482	119	0.476	209	0.552
17	213	0.453	119	0.490	118	0.478	115	0.568
18	113	0.457	114	0.505	111	0.485	118	0.573
19	111	0.463	107	0.508	115	0.510	111	0.588
20	108	0.464	108	0.508	108	0.519	102	0.592
21	116	0.472	113	0.508	121	0.531	119	0.593
22	115	0.485	102	0.527	104	0.551	107	0.603
23	119	0.494	115	0.527	107	0.556	116	0.606
24	121	0.507	121	0.536	134	0.559	123	0.613
25	134	0.508	116	0.540	123	0.563	207	0.614
26	203	0.508	134	0.551	203	0.583	108	0.617
27	123	0.516	123	0.559	116	0.606	203	0.641
28	207	0.519	203	0.567	103	0.607	103	0.644
29	103	0.532	103	0.577	109	0.610	135	0.646
30	104	0.533	104	0.578	201	0.616	109	0.664
31	201	0.542	207	0.589	211	0.617	134	0.669
32	135	0.543	135	0.589	122	0.641	112	0.686
33	102	0.568	201	0.604	120	0.652	121	0.694
34	109	0.569	109	0.609	113	0.654	133	0.698
35	122	0.570	122	0.613	135	0.656	201	0.705
36	112	0.578	211	0.627	207	0.672	104	0.715
37	133	0.579	133	0.638	112	0.686	122	0.745
38	211	0.597	112	0.641	206	0.689	211	0.753
39	120	0.629	120	0.656	133	0.702	206	0.764
40	206	0.648	132	0.676	114	0.726	120	0.802
41	132	0.675	206	0.685	132	0.802	132	0.851
42	214	0.681	214	0.765	214	0.846	214	0.889
43	129	0.883	129	0.882	129	0.902	129	0.946
44	117	0.927	130	0.950	130	0.929	117	1.019
45	130	0.945	117	0.967	212	1.001	212	1.037
46	212	0.974	212	0.988	117	1.038	130	1.101

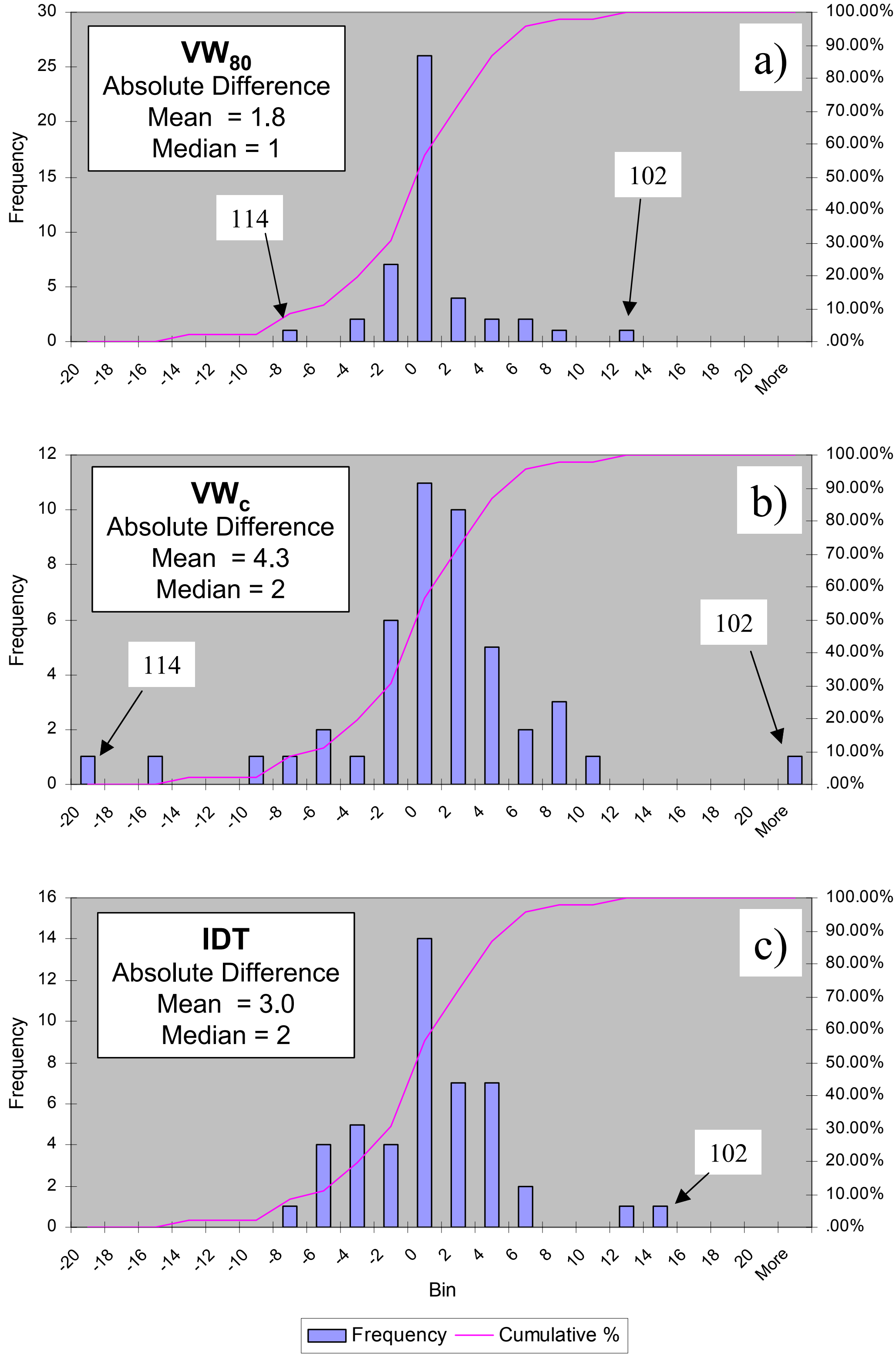


Figure 2-12. Histograms of the Changes in OSF Rankings that Result from Subtracting the Area-Based MOE Rankings from the Nominal MOE Rankings for a Threshold of 0.01 ng m^{-3}: a) VW$_{80}$, b) VW$_{c}$, and c) IDT

associated with all of the area-based MOE values, led to the improved ranking status. At the later time periods, model 102 achieved relatively unbiased (near the diagonal) and accurate predictions, relative to the close-in time and range period where there was a large under-prediction.

Model 114 – CNR (Italy) from Table 1-1 – was ranked 10 using OSF and the nominal MOE and fell to 18 and 40 based on VW_{80} and VW_c MOE values, respectively, but *rose* to 6 based on the IDT technique. This is unusual behavior as model relative ranking performance was generally found to be fairly insensitive to the three area-based techniques. This result is worthy of further discussion.

While model 114's MOE appears to improve with time initially (moving from a large over-prediction toward the diagonal and about the right size at a threshold of 0.01 ng m^{-3}), at the longest times – after 60 hours – there were no predictions provided by this model. Therefore, comparisons with observations were not included for model 114 past the 60 hour mark. Of course, the other model MOEs, in general, were based on both the shorter time/range *and* longer time/range predictions, especially when considering the 0.01 ng m^{-3}. This leads to an artificial comparative advantage for 114, since the longer time/distance predictions will generally correspond to a more difficult prediction problem. The IDT technique, with the initial logarithmic transformation, tends to downplay the importance of the highest concentrations comparisons (at close range and short times) and with the exclusion of the longer time/distance comparisons, the relative advantage for model 114 was substantial and is reflected in its improved ranking. On the other hand, the large over-prediction of model 114 at short times/distances combined with the VW weighting techniques (which did not involve logarithmic transformations) allowed for a more severe relative penalty for model 114 even with the exclusion of the post-60 hour comparisons. Similar behavior for model 114 was observed at the 0.1 ng min m^{-3} threshold. Also, model 113 (ANPA, Italy) only provided predictions up to 60 hours after the release (and this was not true for any other models). As seen in Tables 2-7 and 2-10, model 113 follows the same trend as 114, that is, decreased or similar relative ranking for MOEs based on VW weighting and increased relative ranking based on IDT. Model 113 (for the 0.01 ng min m^{-3} threshold, Table 2-7) was ranked 18 using OSF and the nominal MOE but decreased to 21 and 34 based on VW_{80} and VW_c MOE values, respectively, but *increased* to 7 based on the IDT technique.

Tables 2-8 and 2-9 present RWFMS(1,1) and RWFMS(5,0.5) rankings for nominal and area-based 0.01 ng m^{-3} threshold MOE values, respectively. As was true for OSF, the RWFMS(1,1) rankings are relatively robust to the different techniques (nominal and area-based) used to compute the MOE values. Of the top 10 RWFMS(1,1) ranked models based on the nominal MOE, 8, 7, and 8 appear in the top 10 based on the VW_{80}, VW_c, and IDT MOE values, respectively. Similarly, of the bottom 10 RWFMS(1,1) ranked models based on the nominal MOE, 9, 8, and 9 appear in the bottom 10 based on the VW_{80}, VW_c, and IDT MOE values, respectively. For RWFMS(5,0.5), similar robustness is evident. Of the top 10 RWFMS(5,0.5) ranked models based on the nominal MOE, 8, 5, and 7 appear in the top 10 based on the VW_{80}, VW_c, and IDT MOE values, respectively. Similarly, of the bottom 10 RWFMS(5,0.5) ranked models based on the nominal MOE, 9, 9, and 9 appear in the bottom 10 based on the VW_{80}, VW_c, and IDT MOE values, respectively.

Next, Tables 2-10 through 2-15 present OSF, RWFMS(1,1), and RWFMS(5,0.5) rankings for nominal and area-based MOE values at higher thresholds. Tables 2-10 through 2-12 are based on a threshold of 0.1 ng m^{-3} and Tables 2-13 through 2-15 are based on a threshold of 0.5 ng m^{-3}. Finally, Table 2-16 summarizes the mean and median absolute differences between the various area-based MOE value rankings and the nominal MOE value rankings for all three thresholds and for all three scoring functions. The median and mean absolute ranking differences reported in Table 2-16 suggest that the least differences between the nominal rankings result from the VW_{80} and IDT area-based computations. The VW_c area-based MOE values led to larger differences in rankings when compared to the nominal rankings.

Table 2-8. Comparisons of Model Rankings Based on RWFMS(1,1) for a 0.01 ng m^{-3} Threshold: Nominal and Area-Based – VW_{80}, VW_c, and IDT – MOE Values

Rank	Model	Nominal	Model	VW_{80}	Model	VW_c	Model	IDT
1	202	0.597	105	0.593	105	0.642	105	0.594
2	105	0.594	202	0.579	202	0.598	202	0.546
3	208	0.574	131	0.560	131	0.577	208	0.520
4	127	0.568	208	0.546	127	0.565	127	0.508
5	128	0.565	127	0.542	208	0.559	106	0.499
6	210	0.548	210	0.529	210	0.551	114	0.486
7	131	0.546	128	0.525	102	0.547	113	0.482
8	101	0.545	205	0.520	213	0.539	101	0.469
9	205	0.544	204	0.519	101	0.527	210	0.469
10	114	0.541	213	0.516	204	0.527	128	0.466
11	106	0.537	106	0.515	106	0.516	131	0.464
12	110	0.535	110	0.514	110	0.511	204	0.464
13	118	0.530	101	0.513	205	0.510	205	0.457
14	204	0.526	118	0.503	209	0.508	110	0.445
15	209	0.521	209	0.495	128	0.498	213	0.442
16	107	0.517	111	0.492	119	0.497	209	0.438
17	213	0.516	119	0.487	118	0.496	115	0.422
18	113	0.514	114	0.473	111	0.489	118	0.414
19	111	0.507	107	0.472	115	0.469	111	0.413
20	108	0.506	108	0.471	108	0.461	102	0.408
21	116	0.500	113	0.470	121	0.454	107	0.402
22	115	0.489	102	0.455	104	0.437	116	0.399
23	119	0.485	115	0.454	107	0.435	119	0.398
24	121	0.472	121	0.451	134	0.432	207	0.389
25	134	0.471	116	0.446	123	0.423	108	0.388
26	203	0.470	134	0.437	203	0.415	123	0.382
27	123	0.464	123	0.426	116	0.398	203	0.372
28	207	0.461	203	0.425	103	0.393	135	0.362
29	104	0.452	104	0.417	201	0.392	103	0.361
30	103	0.450	103	0.412	109	0.391	109	0.353
31	201	0.445	207	0.406	211	0.380	134	0.351
32	135	0.441	135	0.403	122	0.362	121	0.342
33	109	0.423	201	0.400	135	0.357	201	0.329
34	122	0.422	122	0.390	120	0.356	104	0.321
35	102	0.419	109	0.390	113	0.350	112	0.316
36	112	0.412	211	0.376	207	0.347	133	0.310
37	133	0.409	133	0.362	206	0.333	122	0.297
38	211	0.397	112	0.362	112	0.332	211	0.292
39	120	0.374	120	0.355	133	0.317	206	0.282
40	206	0.363	132	0.350	114	0.299	120	0.249
41	132	0.344	206	0.334	132	0.276	132	0.244
42	214	0.329	214	0.276	214	0.235	214	0.202
43	130	0.188	130	0.193	130	0.206	130	0.120
44	129	0.120	129	0.124	129	0.110	129	0.085
45	117	0.097	117	0.079	117	0.050	117	0.035
46	212	0.072	212	0.056	212	0.043	212	0.025

Table 2-9. Comparisons of Model Rankings Based on RWFMS(5,0.5) for a 0.01 ng m^{-3} Threshold: Nominal and Area-Based – VW_{80}, VW_c, and IDT – MOE Values

Rank	Model	Nominal	Model	VW_{80}	Model	VW_c	Model	IDT
1	113	0.402	205	0.408	205	0.446	205	0.421
2	101	0.401	110	0.399	101	0.434	113	0.413
3	114	0.396	101	0.394	123	0.396	114	0.400
4	123	0.394	113	0.370	105	0.391	101	0.390
5	135	0.388	123	0.368	110	0.385	110	0.363
6	103	0.384	114	0.346	127	0.377	105	0.346
7	110	0.384	103	0.340	115	0.376	106	0.343
8	205	0.381	115	0.334	202	0.343	123	0.324
9	207	0.355	105	0.330	106	0.342	112	0.314
10	115	0.348	106	0.328	131	0.330	115	0.311
11	116	0.338	135	0.325	204	0.326	127	0.302
12	106	0.334	127	0.323	210	0.295	202	0.282
13	112	0.327	202	0.312	104	0.288	103	0.277
14	127	0.325	204	0.301	213	0.287	135	0.262
15	105	0.322	207	0.290	113	0.286	207	0.261
16	202	0.316	112	0.290	209	0.283	204	0.258
17	203	0.314	109	0.277	102	0.274	109	0.228
18	204	0.289	116	0.274	208	0.271	203	0.222
19	109	0.281	131	0.271	111	0.267	208	0.222
20	104	0.277	203	0.270	109	0.266	116	0.215
21	107	0.276	210	0.268	103	0.265	210	0.214
22	210	0.274	209	0.262	121	0.251	131	0.212
23	214	0.270	104	0.260	112	0.243	209	0.209
24	209	0.269	111	0.253	135	0.241	111	0.203
25	208	0.268	208	0.251	203	0.238	104	0.196
26	128	0.263	107	0.249	114	0.234	107	0.196
27	201	0.254	121	0.240	116	0.229	201	0.194
28	206	0.249	128	0.239	206	0.229	206	0.193
29	121	0.240	213	0.238	128	0.227	128	0.189
30	131	0.240	206	0.236	207	0.225	213	0.187
31	108	0.239	201	0.234	107	0.220	102	0.171
32	111	0.238	108	0.206	201	0.220	121	0.162
33	213	0.222	214	0.204	119	0.205	108	0.150
34	118	0.217	118	0.204	118	0.201	214	0.143
35	134	0.192	119	0.188	134	0.187	118	0.141
36	119	0.179	102	0.186	108	0.178	119	0.136
37	122	0.167	134	0.182	214	0.150	134	0.132
38	102	0.149	122	0.153	211	0.130	122	0.100
39	211	0.140	132	0.139	122	0.122	211	0.098
40	133	0.138	211	0.134	120	0.122	133	0.092
41	120	0.133	120	0.126	132	0.121	132	0.089
42	132	0.123	133	0.120	133	0.103	120	0.075
43	130	0.061	130	0.069	130	0.081	130	0.039
44	129	0.027	129	0.028	129	0.025	129	0.019
45	117	0.022	117	0.018	117	0.011	117	0.007
46	212	0.016	212	0.012	212	0.009	212	0.005

Table 2-10. Comparisons of Model Rankings Based on OSF for a 0.1 ng m^{-3} Threshold: Nominal and Area-Based – VW_{80}, VW_c, and IDT – MOE Values

Rank	Model	Nominal	Model	VW_{80}	Model	VW_c	Model	IDT
1	208	0.381	208	0.426	107	0.502	208	0.442
2	128	0.411	202	0.462	101	0.502	105	0.468
3	202	0.419	101	0.468	127	0.515	202	0.474
4	101	0.424	128	0.481	202	0.515	128	0.493
5	127	0.440	131	0.482	105	0.519	127	0.501
6	107	0.446	105	0.489	131	0.520	114	0.540
7	105	0.451	127	0.493	115	0.548	101	0.544
8	131	0.462	107	0.519	208	0.563	210	0.569
9	118	0.476	115	0.523	128	0.567	209	0.573
10	115	0.481	111	0.528	213	0.570	106	0.578
11	205	0.488	118	0.530	134	0.576	113	0.580
12	134	0.492	210	0.535	204	0.580	131	0.588
13	106	0.494	106	0.537	111	0.582	115	0.595
14	210	0.495	213	0.544	209	0.582	118	0.611
15	114	0.499	205	0.547	102	0.584	107	0.613
16	111	0.505	209	0.549	118	0.598	205	0.618
17	209	0.509	204	0.553	106	0.599	213	0.618
18	204	0.513	114	0.560	211	0.606	111	0.623
19	213	0.522	134	0.560	203	0.610	119	0.635
20	110	0.526	119	0.569	210	0.630	110	0.645
21	133	0.526	110	0.571	123	0.644	207	0.658
22	119	0.547	102	0.580	104	0.655	133	0.661
23	113	0.560	113	0.604	121	0.666	204	0.666
24	207	0.562	123	0.608	108	0.673	123	0.667
25	123	0.565	207	0.620	110	0.676	134	0.687
26	102	0.572	133	0.643	119	0.698	102	0.710
27	201	0.594	203	0.651	120	0.699	201	0.717
28	203	0.606	211	0.668	207	0.701	103	0.724
29	211	0.612	104	0.672	122	0.716	203	0.736
30	121	0.637	121	0.674	205	0.722	108	0.741
31	108	0.638	201	0.676	114	0.745	121	0.754
32	104	0.647	122	0.689	113	0.746	122	0.772
33	103	0.652	109	0.717	201	0.793	109	0.774
34	122	0.653	120	0.720	109	0.811	104	0.776
35	120	0.671	103	0.720	133	0.835	135	0.783
36	135	0.687	108	0.724	103	0.836	116	0.787
37	116	0.694	206	0.808	206	0.846	211	0.799
38	109	0.695	112	0.816	112	0.858	206	0.842
39	112	0.741	129	0.824	214	0.891	120	0.851
40	206	0.778	116	0.833	129	0.899	112	0.855
41	132	0.803	132	0.842	132	0.919	214	0.951
42	214	0.817	135	0.852	116	0.949	129	0.974
43	129	0.822	214	0.882	117	1.045	132	0.980
44	117	0.927	117	0.974	212	1.100	117	1.074
45	212	1.071	212	1.093	135	1.110	212	1.110
46	130	1.092	130	1.102	130	1.153	130	1.222

Table 2-11. Comparisons of Model Rankings Based on RWFMS(1,1) for a 0.1 ng m^{-3} Threshold: Nominal and Area-Based – VW_{80}, VW_c, and IDT – MOE Values

Rank	Model	Nominal	Model	VW_{80}	Model	VW_c	Model	IDT
1	208	0.577	208	0.538	107	0.476	208	0.524
2	128	0.551	202	0.508	101	0.476	105	0.503
3	202	0.545	101	0.505	127	0.466	202	0.498
4	101	0.544	128	0.492	202	0.464	128	0.483
5	127	0.526	131	0.492	131	0.462	127	0.476
6	107	0.521	105	0.486	105	0.461	114	0.444
7	105	0.517	127	0.482	115	0.438	101	0.441
8	131	0.508	107	0.463	208	0.430	210	0.425
9	118	0.497	115	0.458	128	0.427	209	0.421
10	115	0.493	111	0.456	213	0.426	131	0.413
11	205	0.487	118	0.452	134	0.421	106	0.411
12	134	0.484	210	0.451	111	0.417	113	0.410
13	106	0.482	106	0.447	204	0.415	115	0.399
14	210	0.481	213	0.445	209	0.414	107	0.394
15	114	0.478	209	0.439	102	0.412	213	0.392
16	111	0.473	205	0.438	211	0.399	111	0.389
17	209	0.470	204	0.436	203	0.397	118	0.384
18	204	0.467	134	0.433	106	0.396	205	0.374
19	213	0.460	114	0.432	118	0.395	119	0.374
20	110	0.456	119	0.424	210	0.383	204	0.355
21	133	0.455	102	0.418	121	0.358	207	0.354
22	119	0.439	110	0.416	104	0.358	110	0.350
23	113	0.428	113	0.396	123	0.353	133	0.348
24	207	0.426	207	0.386	108	0.345	134	0.345
25	102	0.423	123	0.384	119	0.339	102	0.331
26	123	0.421	203	0.369	110	0.338	123	0.330
27	201	0.403	133	0.367	207	0.334	203	0.313
28	203	0.399	211	0.358	120	0.333	201	0.310
29	211	0.396	121	0.354	122	0.318	108	0.310
30	121	0.379	104	0.347	205	0.309	121	0.303
31	108	0.378	201	0.346	114	0.305	122	0.292
32	122	0.368	122	0.344	113	0.294	103	0.290
33	104	0.365	120	0.322	201	0.280	135	0.285
34	120	0.354	109	0.322	109	0.271	116	0.285
35	103	0.351	108	0.319	103	0.248	109	0.280
36	135	0.345	103	0.306	133	0.246	211	0.277
37	116	0.341	112	0.263	112	0.242	104	0.265
38	109	0.336	116	0.256	206	0.232	120	0.236
39	112	0.306	206	0.250	214	0.227	112	0.231
40	132	0.270	132	0.250	132	0.211	206	0.216
41	206	0.269	135	0.248	116	0.195	214	0.186
42	214	0.263	214	0.230	129	0.157	132	0.178
43	129	0.190	129	0.195	135	0.118	129	0.093
44	130	0.129	130	0.123	130	0.099	130	0.073
45	117	0.125	117	0.102	117	0.067	117	0.030
46	212	0.060	212	0.044	212	0.036	212	0.014

Table 2-12. Comparisons of Model Rankings Based on RWFMS(5,0.5) for a 0.1 ng m^{-3} Threshold: Nominal and Area-Based – VW_{80}, VW_c, and IDT – MOE Values

Rank	Model	Nominal	Model	VW_{80}	Model	VW_c	Model	IDT
1	101	0.438	101	0.410	101	0.395	101	0.432
2	110	0.409	110	0.379	105	0.350	205	0.379
3	123	0.385	202	0.350	202	0.340	105	0.373
4	205	0.381	123	0.345	123	0.330	110	0.361
5	208	0.380	205	0.345	115	0.311	123	0.358
6	202	0.377	208	0.338	106	0.298	202	0.350
7	115	0.357	105	0.337	127	0.290	106	0.336
8	106	0.357	115	0.326	131	0.290	114	0.336
9	105	0.350	106	0.325	102	0.271	113	0.326
10	128	0.344	127	0.304	204	0.265	127	0.316
11	127	0.339	128	0.283	209	0.264	115	0.311
12	207	0.323	209	0.278	110	0.251	208	0.293
13	114	0.309	131	0.270	107	0.246	103	0.268
14	113	0.305	113	0.262	104	0.245	209	0.255
15	209	0.301	204	0.261	208	0.241	128	0.251
16	107	0.300	210	0.259	134	0.230	207	0.248
17	103	0.294	114	0.258	128	0.226	210	0.240
18	210	0.282	107	0.245	205	0.215	201	0.223
19	201	0.275	207	0.241	213	0.213	107	0.211
20	204	0.271	103	0.229	210	0.203	204	0.210
21	131	0.265	104	0.229	113	0.202	104	0.204
22	134	0.246	102	0.224	111	0.200	131	0.203
23	104	0.238	213	0.221	121	0.196	111	0.184
24	213	0.227	201	0.220	203	0.186	213	0.180
25	111	0.221	111	0.218	207	0.181	206	0.180
26	203	0.217	134	0.218	211	0.179	109	0.178
27	118	0.207	109	0.187	114	0.172	203	0.165
28	109	0.198	203	0.181	206	0.155	102	0.159
29	102	0.196	119	0.180	119	0.150	121	0.153
30	206	0.193	206	0.179	103	0.142	135	0.145
31	211	0.188	118	0.174	201	0.141	112	0.145
32	121	0.185	121	0.174	118	0.135	134	0.144
33	133	0.183	211	0.167	120	0.125	119	0.141
34	135	0.182	122	0.150	109	0.125	118	0.131
35	112	0.180	112	0.143	108	0.123	116	0.125
36	119	0.176	133	0.137	112	0.120	108	0.122
37	122	0.166	120	0.126	122	0.114	133	0.117
38	108	0.166	108	0.123	214	0.103	122	0.117
39	116	0.154	214	0.111	133	0.083	211	0.115
40	120	0.145	135	0.103	132	0.081	214	0.104
41	214	0.140	116	0.099	116	0.074	120	0.078
42	132	0.099	132	0.093	130	0.046	132	0.063
43	130	0.051	130	0.054	135	0.040	130	0.027
44	129	0.047	129	0.049	129	0.039	129	0.021
45	117	0.030	117	0.024	117	0.015	117	0.006
46	212	0.014	212	0.010	212	0.008	212	0.003

Table 2-13. Comparisons of Model Rankings Based on OSF for a 0.5 ng m^{-3} Threshold: Nominal and Area-Based – VW_{80}, VW_c, and IDT – MOE Values

Rank	Model	Nominal	Model	VW_{80}	Model	VW_c	Model	IDT
1	127	0.600	127	0.627	127	0.633	127	0.724
2	107	0.632	107	0.688	118	0.709	107	0.724
3	134	0.635	118	0.701	205	0.811	128	0.728
4	118	0.646	128	0.707	113	0.822	111	0.756
5	128	0.658	111	0.708	209	0.823	134	0.780
6	208	0.676	134	0.720	128	0.834	208	0.784
7	111	0.676	208	0.728	208	0.834	105	0.793
8	133	0.682	209	0.736	107	0.859	106	0.818
9	131	0.687	205	0.736	210	0.868	205	0.824
10	205	0.704	131	0.747	134	0.868	110	0.842
11	209	0.709	105	0.764	110	0.870	209	0.851
12	105	0.714	113	0.771	131	0.873	133	0.858
13	119	0.724	110	0.781	111	0.874	118	0.859
14	110	0.735	133	0.790	105	0.894	202	0.882
15	101	0.746	202	0.796	101	0.900	210	0.918
16	210	0.750	210	0.798	213	0.908	131	0.929
17	202	0.750	123	0.820	123	0.909	113	0.937
18	113	0.752	106	0.826	202	0.922	213	0.944
19	106	0.756	101	0.829	115	0.957	101	0.945
20	213	0.766	213	0.831	133	0.958	207	0.946
21	123	0.804	119	0.841	106	0.961	204	0.951
22	207	0.807	204	0.878	104	0.968	201	0.953
23	121	0.822	203	0.879	211	0.969	123	0.959
24	204	0.832	207	0.908	201	0.969	119	0.967
25	203	0.841	201	0.908	103	0.972	121	0.982
26	201	0.846	115	0.909	122	0.974	203	1.005
27	102	0.870	104	0.911	204	0.989	104	1.019
28	104	0.874	121	0.919	119	0.998	211	1.021
29	115	0.876	122	0.924	121	1.000	115	1.029
30	116	0.879	102	0.930	108	1.011	120	1.042
31	211	0.920	103	0.939	207	1.021	103	1.062
32	120	0.922	211	0.962	120	1.024	108	1.086
33	122	0.928	116	0.968	203	1.025	206	1.088
34	132	0.948	132	0.972	102	1.046	102	1.102
35	103	0.954	120	0.984	112	1.054	116	1.103
36	108	0.977	108	1.010	132	1.070	122	1.125
37	206	0.993	206	1.034	206	1.079	114	1.167
38	112	0.995	114	1.037	116	1.102	112	1.193
39	114	1.002	112	1.042	114	1.103	135	1.211
40	129	1.025	129	1.080	109	1.160	132	1.214
41	214	1.028	135	1.104	129	1.175	214	1.225
42	135	1.052	214	1.112	212	1.191	129	1.242
43	117	1.122	117	1.152	117	1.200	212	1.271
44	109	1.156	109	1.159	214	1.205	109	1.287
45	212	1.173	130	1.186	135	1.207	117	1.312
46	130	1.210	212	1.191	130	1.233	130	1.358

Table 2-14. Comparisons of Model Rankings Based on RWFMS(1,1) for a 0.5 ng m^{-3} Threshold: Nominal and Area-Based – VW_{80}, VW_c, and IDT – MOE Values

Rank	Model	Nominal	Model	VW_{80}	Model	VW_c	Model	IDT
1	127	0.401	127	0.385	127	0.380	107	0.313
2	107	0.381	107	0.344	118	0.321	127	0.310
3	134	0.371	111	0.331	113	0.260	111	0.303
4	118	0.368	118	0.325	209	0.242	128	0.277
5	111	0.352	134	0.310	205	0.234	134	0.263
6	133	0.349	209	0.308	210	0.232	208	0.252
7	128	0.344	128	0.307	111	0.228	118	0.237
8	131	0.341	208	0.304	107	0.223	133	0.237
9	208	0.338	131	0.300	208	0.216	209	0.226
10	209	0.325	205	0.296	101	0.216	205	0.226
11	119	0.323	113	0.293	131	0.203	105	0.224
12	205	0.319	133	0.282	128	0.194	210	0.203
13	101	0.309	210	0.277	213	0.194	131	0.195
14	210	0.305	101	0.260	115	0.191	106	0.193
15	113	0.300	105	0.257	134	0.186	101	0.192
16	105	0.299	119	0.253	211	0.177	119	0.187
17	213	0.290	213	0.251	123	0.171	204	0.187
18	110	0.278	123	0.244	133	0.165	213	0.186
19	202	0.272	110	0.241	103	0.163	110	0.184
20	106	0.267	203	0.232	121	0.163	113	0.178
21	121	0.265	202	0.232	108	0.161	121	0.178
22	207	0.264	204	0.230	110	0.159	202	0.169
23	123	0.256	115	0.217	122	0.159	203	0.168
24	204	0.255	207	0.212	120	0.153	207	0.161
25	203	0.253	121	0.212	204	0.152	115	0.151
26	115	0.235	106	0.211	203	0.150	201	0.146
27	116	0.232	201	0.198	201	0.150	120	0.145
28	201	0.231	122	0.196	112	0.144	123	0.143
29	102	0.230	116	0.187	119	0.144	211	0.142
30	104	0.212	103	0.187	104	0.139	108	0.131
31	120	0.207	104	0.186	105	0.137	116	0.120
32	211	0.201	102	0.185	132	0.136	104	0.108
33	122	0.198	132	0.183	207	0.130	103	0.106
34	132	0.196	211	0.182	114	0.124	102	0.106
35	108	0.183	120	0.177	202	0.115	122	0.099
36	103	0.180	108	0.162	116	0.111	114	0.095
37	112	0.170	114	0.151	109	0.094	112	0.084
38	114	0.168	112	0.148	206	0.091	206	0.081
39	214	0.157	129	0.134	102	0.086	135	0.076
40	129	0.155	206	0.129	106	0.086	132	0.074
41	206	0.148	135	0.122	129	0.084	214	0.069
42	135	0.143	214	0.120	117	0.080	129	0.059
43	117	0.103	109	0.094	135	0.070	109	0.047
44	109	0.099	117	0.093	214	0.065	130	0.018
45	130	0.072	130	0.072	212	0.044	117	0.016
46	212	0.059	212	0.044	130	0.032	212	0.009

Table 2-15. Comparisons of Model Rankings Based on RWFMS(5,0.5) for a 0.5 ng m^{-3} Threshold: Nominal and Area-Based – VW_{80}, VW_c, and IDT – MOE Values

Rank	Model	Nominal	Model	VW_{80}	Model	VW_c	Model	IDT
1	128	0.301	128	0.266	128	0.213	128	0.303
2	110	0.273	105	0.247	127	0.207	105	0.246
3	105	0.270	110	0.241	205	0.198	106	0.244
4	134	0.261	202	0.227	110	0.192	208	0.213
5	208	0.256	134	0.219	208	0.189	110	0.213
6	127	0.246	205	0.216	134	0.177	127	0.207
7	202	0.245	208	0.215	105	0.175	134	0.202
8	106	0.242	127	0.209	209	0.170	107	0.198
9	205	0.235	106	0.206	131	0.158	205	0.191
10	107	0.211	209	0.184	202	0.152	202	0.174
11	131	0.200	131	0.181	107	0.152	209	0.156
12	209	0.198	107	0.174	123	0.148	111	0.133
13	133	0.177	123	0.172	113	0.140	133	0.132
14	123	0.175	113	0.145	213	0.133	201	0.128
15	213	0.169	213	0.145	210	0.128	207	0.126
16	113	0.169	111	0.137	111	0.126	123	0.126
17	210	0.159	104	0.136	104	0.121	113	0.123
18	201	0.156	133	0.136	106	0.120	210	0.115
19	207	0.155	210	0.135	133	0.116	131	0.114
20	111	0.153	201	0.129	201	0.116	213	0.109
21	104	0.147	102	0.122	101	0.115	204	0.103
22	101	0.146	119	0.122	118	0.111	101	0.102
23	118	0.143	101	0.122	122	0.109	104	0.100
24	119	0.142	122	0.117	103	0.107	211	0.090
25	204	0.132	204	0.117	204	0.103	121	0.085
26	102	0.128	103	0.113	119	0.102	118	0.083
27	211	0.115	118	0.112	211	0.098	119	0.083
28	121	0.114	207	0.112	207	0.095	103	0.083
29	122	0.111	211	0.099	115	0.090	115	0.077
30	103	0.107	121	0.094	102	0.089	203	0.076
31	120	0.103	115	0.094	121	0.087	206	0.075
32	206	0.103	203	0.089	203	0.081	120	0.074
33	115	0.102	206	0.089	120	0.080	102	0.067
34	203	0.102	120	0.083	206	0.078	122	0.059
35	116	0.090	116	0.073	116	0.064	116	0.057
36	132	0.073	132	0.068	132	0.063	108	0.052
37	108	0.071	108	0.055	108	0.055	114	0.040
38	114	0.060	114	0.053	114	0.051	112	0.036
39	112	0.060	112	0.052	112	0.051	135	0.033
40	214	0.057	129	0.050	129	0.045	214	0.032
41	129	0.052	214	0.048	214	0.041	132	0.024
42	135	0.049	130	0.045	135	0.039	109	0.018
43	130	0.037	135	0.044	130	0.032	129	0.018
44	109	0.034	109	0.030	109	0.030	130	0.009
45	117	0.030	117	0.028	117	0.027	117	0.004
46	212	0.014	212	0.010	212	0.010	212	0.002

Table 2-16. Median / Mean Absolute Ranking Difference Between the Area-Based MOE Value Rankings and Nominal MOE Value Rankings for 3 Thresholds and for 3 Scoring Functions

($ng\ m^{-3}$) →	0.01	0.1	0.5	0.01	0.1	0.5	0.01	0.1	0.5
Scoring Function	VW_{80}			VW_c			IDT		
OSF	1/1.8	2/2.2	1/1.9	2/4.3	4.5/5.5	3.5/4.6	2/3.0	2/3.1	2/2.9
RWFMS(1,1)	1/1.9	2/2.0	1/2.1	2/4.1	4.5/5.4	5.5/6.7	1.5/2.7	2/3.0	2/2.7
RWFMS(5,0.5)	2/2.9	2/2.2	1.5/1.8	7/7.3	5/5.7	2/2.3	2/3.2	2/2.6	2/2.4

Next we considered the top 10 and bottom 10 ranked models based on the OSF, RWFMS(1,1), and RWFMS(5,0.5) scoring functions at each of the examined thresholds. Table 2-17 identifies the number of models that were in the top 10 (or bottom 10) based on the nominal MOE values *and* the area-based MOE values for the three techniques: VW_{80}, VW_c, and IDT.

Table 2-17. Agreements of the Top 10 and Bottom 10 Ranked Models for 3 Different Thresholds and 3 Different Scoring Functions When Compared to the Rankings Based on the Nominal MOE Values

Agreements for Top and Bottom Ten for 0.01 ng m^{-3} Threshold			
	VW_{80}	VWc	IDT
OSF	8 / 9	7 / 8	8 / 9
RWFMS(1,1)	8 / 9	7 / 8	8 / 9
RWFMS(5,0.5)	8 / 9	5 / 9	7 / 9

Agreements for Top and Bottom Ten for 0.1 ng m^{-3} Threshold			
	VW_{80}	VWc	IDT
OSF	9 / 9	9 / 9	6 / 8
RWFMS(1,1)	9 / 9	9 / 9	7 / 8
RWFMS(5,0.5)	9 / 9	6 / 8	7 / 8

Agreements for Top and Bottom Ten for 0.5 ng m^{-3} Threshold			
	VW_{80}	VWc	IDT
OSF	9 / 10	7 / 9	7 / 9
RWFMS(1,1)	9 / 10	6 / 8	9 / 9
RWFMS(5,0.5)	9 / 10	8 / 10	10 / 9

Overall, the rankings based on VW_{80} are most like those based on the nominal MOE values. Of course, some of the differences associated with the IDT based rankings should reflect real changes that are associated with various models being relatively better or worse at predicting the longer range patterns. Of course, at these longer ranges where the sampler density is lower and hence the area being "represented" is larger, the IDT technique (by design) weights these results as more important.

Tables 2-18 through 2-21 compare rankings based on summed concentration MOE values (as opposed to threshold based) for the nominal, VW_{80}, VW_c, and IDT cases. For the four tables the scoring functions OSF, RWFMS(1,1), RWFMS(5,0.5,) and absolute fractional bias – ABS(FB)[9] – are considered. Figure 2-13 shows histograms of the changes in OSF rankings that result from subtracting the three types of area-based MOE rankings from the nominal MOE rankings. The biggest changes are associated with very large improved rankings for models 127, 118, and 121 when the IDT area-based technique is used. These three models were ranked (by OSF) as 33, 20, and 41, respectively, based on the nominal MOE. After applying the IDT area technique for the computation of the MOE, models 127, 118, and 121 are ranked as 1, 6, and 28, respectively. These improved rankings mirror changes seen for these same three models in the previous study when the single "near-release" sampler at Rennes was removed (Ref. 2-1). As reported in Ref. 2-1, the OSF-based rankings of models 127 (ARAC), 118 (FOA), and 121 (SCIPUFF) were 8, 4, and 34, respectively, after the removal of the single Rennes sampler location. Hence, the sensitivity of model rankings to this single sampler location were previously described and the IDT technique (which involves first a logarithmic transformation followed by linear interpolation) appears to mitigate the dominating effects of this single sampler location.

Finally, Table 2-22 identifies the number of top 10 and bottom 10 ranked models that are in agreement with those based on the nominal MOE values for the three area-based techniques: VW_{80}, VW_c, and IDT.

---

9 $FB = \dfrac{\overline{C_p} - \overline{C_o}}{0.5\left(\overline{C_o} + \overline{C_p}\right)}$, where C = observation/prediction of interest (e.g., dosage), C_p corresponds to model prediction, C_o corresponds to observation, and $\overline{C}$ denotes the average. FB has previously been related to the x and y coordinates of the MOE as follows: $FB = \dfrac{2(x-y)}{x+y}$. See Reference 2-1.

Table 2-18. Comparisons of Model Rankings Based on OSF for Summed Concentration: Nominal and Area-Based – VW_{80}, VW_c, and IDT – MOE Values

Rank	Model	Nominal	Model	VW_{80}	Model	VW_c	Model	IDT
1	107	0.625	107	0.657	101	0.663	127	0.673
2	205	0.669	205	0.672	107	0.697	111	0.682
3	110	0.689	101	0.692	115	0.733	107	0.684
4	101	0.690	110	0.709	205	0.735	101	0.684
5	113	0.733	113	0.751	209	0.745	205	0.720
6	115	0.744	209	0.753	111	0.753	118	0.723
7	209	0.754	111	0.756	210	0.761	210	0.736
8	123	0.760	123	0.756	213	0.764	209	0.738
9	114	0.760	115	0.757	131	0.772	208	0.752
10	111	0.762	131	0.760	110	0.775	110	0.757
11	131	0.762	213	0.773	208	0.776	115	0.762
12	210	0.776	210	0.777	123	0.786	131	0.780
13	213	0.778	208	0.779	105	0.798	113	0.784
14	203	0.786	202	0.790	113	0.801	204	0.790
15	208	0.787	114	0.795	204	0.803	123	0.799
16	202	0.800	105	0.797	202	0.803	105	0.807
17	128	0.810	203	0.803	203	0.829	202	0.814
18	105	0.812	204	0.807	128	0.842	114	0.820
19	204	0.815	128	0.811	104	0.859	203	0.830
20	118	0.852	103	0.854	108	0.874	119	0.831
21	103	0.854	119	0.866	118	0.881	128	0.834
22	119	0.857	118	0.883	211	0.890	213	0.841
23	135	0.888	104	0.896	103	0.890	103	0.870
24	207	0.890	201	0.918	114	0.901	201	0.875
25	108	0.894	106	0.921	119	0.906	106	0.876
26	104	0.896	108	0.925	127	0.911	108	0.876
27	112	0.900	206	0.928	201	0.925	207	0.914
28	106	0.907	112	0.942	109	0.926	121	0.918
29	201	0.908	135	0.945	206	0.947	104	0.925
30	134	0.914	207	0.945	102	0.948	112	0.941
31	206	0.926	109	0.951	112	0.952	109	0.941
32	214	0.946	211	0.957	106	0.957	211	0.954
33	127	0.954	127	0.960	120	1.001	134	0.969
34	211	0.964	134	0.964	207	1.012	206	0.970
35	109	0.976	102	0.986	134	1.033	135	0.975
36	132	0.979	132	0.988	132	1.035	102	1.009
37	116	0.986	214	1.013	121	1.051	116	1.034
38	133	1.001	120	1.031	135	1.078	120	1.052
39	102	1.002	116	1.037	122	1.086	133	1.055
40	120	1.027	133	1.063	116	1.107	132	1.081
41	121	1.083	122	1.083	214	1.139	214	1.092
42	122	1.096	121	1.084	133	1.142	122	1.105
43	129	1.158	129	1.168	130	1.218	117	1.252
44	117	1.217	130	1.207	129	1.225	129	1.265
45	130	1.225	117	1.239	117	1.260	212	1.268
46	212	1.300	212	1.305	212	1.307	130	1.299

Table 2-19. Comparisons of Model Rankings Based on RWFMS(1,1) for Summed Concentration: Nominal and Area-Based – VW_{80}, VW_c, and IDT – MOE Values

Rank	Model	Nominal	Model	VW_{80}	Model	VW_c	Model	IDT
1	107	0.387	107	0.365	101	0.356	111	0.349
2	205	0.347	205	0.345	107	0.338	107	0.345
3	101	0.343	101	0.340	115	0.314	127	0.343
4	110	0.324	110	0.308	111	0.305	101	0.334
5	113	0.314	113	0.304	205	0.299	118	0.312
6	115	0.310	111	0.302	209	0.295	210	0.311
7	114	0.300	115	0.300	210	0.294	209	0.302
8	111	0.299	123	0.290	213	0.289	205	0.293
9	123	0.289	209	0.288	131	0.274	115	0.288
10	209	0.285	131	0.286	204	0.272	131	0.282
11	203	0.285	213	0.285	113	0.271	208	0.282
12	131	0.284	210	0.280	208	0.267	204	0.277
13	213	0.281	114	0.280	203	0.261	114	0.261
14	210	0.280	203	0.275	110	0.259	113	0.260
15	204	0.269	204	0.273	123	0.257	119	0.260
16	208	0.249	208	0.258	118	0.232	203	0.258
17	118	0.239	202	0.238	108	0.227	110	0.254
18	103	0.237	103	0.238	211	0.227	213	0.246
19	119	0.236	105	0.229	202	0.225	108	0.233
20	202	0.229	119	0.228	105	0.225	123	0.230
21	135	0.228	118	0.224	104	0.225	202	0.219
22	108	0.224	128	0.222	103	0.220	105	0.213
23	112	0.221	104	0.207	114	0.219	128	0.203
24	128	0.214	108	0.204	128	0.212	121	0.202
25	105	0.214	112	0.200	109	0.209	103	0.198
26	104	0.207	135	0.198	119	0.201	109	0.195
27	214	0.196	109	0.196	112	0.196	201	0.193
28	206	0.186	211	0.191	201	0.193	112	0.191
29	211	0.186	201	0.188	206	0.173	211	0.189
30	207	0.186	206	0.186	120	0.167	135	0.176
31	201	0.185	132	0.171	127	0.152	104	0.160
32	109	0.183	214	0.158	132	0.146	207	0.159
33	132	0.176	207	0.157	102	0.142	106	0.145
34	116	0.151	120	0.148	135	0.131	120	0.142
35	134	0.149	102	0.141	207	0.118	102	0.138
36	102	0.145	116	0.126	106	0.095	206	0.133
37	120	0.142	106	0.125	122	0.092	134	0.133
38	106	0.133	134	0.114	116	0.087	132	0.131
39	127	0.115	127	0.113	214	0.077	116	0.131
40	133	0.097	129	0.090	121	0.068	214	0.115
41	129	0.096	122	0.077	129	0.061	133	0.093
42	117	0.073	130	0.072	134	0.057	122	0.087
43	130	0.069	133	0.066	130	0.055	129	0.055
44	122	0.067	117	0.064	117	0.054	130	0.040
45	121	0.047	121	0.047	133	0.028	117	0.024
46	212	0.036	212	0.028	212	0.026	212	0.011

Table 2-20. Comparisons of Model Rankings Based on RWFMS(5,0.5) for Summed Concentration: Nominal and Area-Based – VW_{80}, VW_c, and IDT – MOE Values

Rank	Model	Nominal	Model	VW_{80}	Model	VW_c	Model	IDT
1	110	0.264	110	0.253	105	0.234	110	0.274
2	205	0.238	205	0.237	202	0.224	205	0.269
3	128	0.229	105	0.230	110	0.219	127	0.242
4	105	0.225	202	0.228	101	0.216	101	0.240
5	202	0.224	128	0.216	205	0.211	105	0.237
6	208	0.215	208	0.213	123	0.203	123	0.224
7	209	0.205	209	0.201	208	0.201	208	0.217
8	107	0.195	131	0.191	209	0.199	202	0.215
9	131	0.190	123	0.190	131	0.194	128	0.204
10	123	0.184	101	0.189	128	0.183	113	0.200
11	101	0.181	107	0.174	213	0.176	209	0.199
12	210	0.174	210	0.172	107	0.172	107	0.193
13	113	0.169	213	0.167	115	0.172	106	0.193
14	213	0.169	113	0.157	210	0.168	115	0.183
15	106	0.164	115	0.156	127	0.156	210	0.174
16	207	0.155	106	0.154	113	0.151	103	0.165
17	134	0.155	119	0.137	104	0.149	201	0.164
18	115	0.152	103	0.136	204	0.144	131	0.162
19	201	0.140	104	0.131	111	0.136	111	0.158
20	118	0.138	201	0.129	102	0.133	204	0.156
21	103	0.137	127	0.129	119	0.127	207	0.151
22	119	0.135	207	0.129	106	0.126	104	0.142
23	127	0.132	134	0.127	103	0.120	213	0.139
24	104	0.131	204	0.126	201	0.119	114	0.139
25	111	0.126	111	0.124	206	0.118	203	0.129
26	206	0.125	206	0.123	203	0.117	134	0.122
27	204	0.124	118	0.123	114	0.106	206	0.121
28	114	0.124	114	0.120	118	0.105	121	0.116
29	203	0.123	203	0.114	207	0.102	119	0.110
30	133	0.107	102	0.110	211	0.102	118	0.108
31	116	0.105	135	0.091	109	0.088	112	0.107
32	135	0.102	116	0.089	112	0.083	109	0.102
33	102	0.099	132	0.088	120	0.081	102	0.098
34	214	0.093	211	0.088	121	0.081	211	0.096
35	211	0.092	109	0.083	134	0.079	135	0.094
36	132	0.091	214	0.083	132	0.078	108	0.089
37	112	0.088	112	0.081	108	0.078	116	0.088
38	120	0.086	120	0.079	122	0.075	133	0.086
39	108	0.086	133	0.077	116	0.068	120	0.070
40	109	0.079	122	0.075	135	0.063	122	0.069
41	122	0.070	108	0.072	214	0.059	214	0.067
42	121	0.063	121	0.062	133	0.040	132	0.045
43	129	0.045	129	0.045	130	0.038	129	0.019
44	130	0.032	130	0.038	129	0.036	130	0.019
45	117	0.025	117	0.021	117	0.017	117	0.005
46	212	0.010	212	0.007	212	0.007	212	0.002

Table 2-21. Comparisons of Model Rankings Based on ABS(FB) for Summed Concentration: Nominal and Area-Based – VW_{80}, VW_c, and IDT – MOE Values

Rank	Model	Nominal	Model	VW_{80}	Model	VW_c	Model	IDT
1	203	0.046	107	0.003	109	0.010	119	0.042
2	107	0.061	109	0.043	111	0.018	111	0.046
3	109	0.083	112	0.048	112	0.031	108	0.201
4	135	0.085	114	0.051	203	0.049	203	0.207
5	204	0.086	204	0.072	211	0.088	129	0.221
6	111	0.105	203	0.095	118	0.103	107	0.237
7	115	0.117	111	0.148	107	0.128	132	0.283
8	112	0.122	135	0.164	114	0.223	210	0.288
9	114	0.134	211	0.178	115	0.254	114	0.293
10	101	0.169	113	0.190	204	0.265	211	0.331
11	108	0.184	115	0.209	120	0.307	204	0.340
12	214	0.226	101	0.235	101	0.311	131	0.355
13	113	0.235	108	0.318	210	0.333	120	0.362
14	211	0.299	117	0.350	113	0.336	109	0.375
15	117	0.302	213	0.377	135	0.355	213	0.392
16	129	0.365	118	0.386	103	0.393	135	0.434
17	132	0.374	132	0.392	213	0.417	130	0.459
18	130	0.406	103	0.422	108	0.421	115	0.463
19	213	0.414	210	0.434	132	0.482	112	0.463
20	119	0.430	214	0.434	117	0.499	118	0.465
21	118	0.431	205	0.454	209	0.516	209	0.481
22	103	0.436	129	0.482	205	0.552	121	0.482
23	210	0.452	120	0.483	201	0.590	127	0.484
24	205	0.453	119	0.494	131	0.599	101	0.517
25	123	0.464	123	0.496	119	0.604	214	0.664
26	131	0.524	131	0.520	104	0.619	208	0.665
27	209	0.595	209	0.564	208	0.671	113	0.700
28	104	0.596	104	0.594	123	0.733	205	0.790
29	110	0.643	130	0.608	206	0.739	116	0.821
30	120	0.659	110	0.675	110	0.786	102	0.855
31	206	0.700	206	0.688	129	0.787	103	0.891
32	212	0.784	201	0.719	128	0.899	201	0.918
33	102	0.809	208	0.768	202	0.979	123	0.957
34	116	0.812	116	0.881	105	1.012	110	0.977
35	208	0.824	202	0.924	130	1.026	202	0.981
36	201	0.824	102	0.952	116	1.083	104	1.029
37	207	0.923	207	0.962	214	1.087	128	1.036
38	202	0.960	128	0.970	122	1.107	105	1.080
39	105	1.045	105	0.981	207	1.109	122	1.090
40	128	1.055	212	1.091	102	1.117	207	1.091
41	134	1.186	134	1.320	212	1.153	206	1.113
42	106	1.340	122	1.321	127	1.163	134	1.125
43	127	1.340	127	1.338	106	1.490	133	1.231
44	133	1.361	106	1.358	121	1.500	106	1.352
45	122	1.389	133	1.490	134	1.620	117	1.565
46	121	1.623	121	1.627	133	1.726	212	1.790

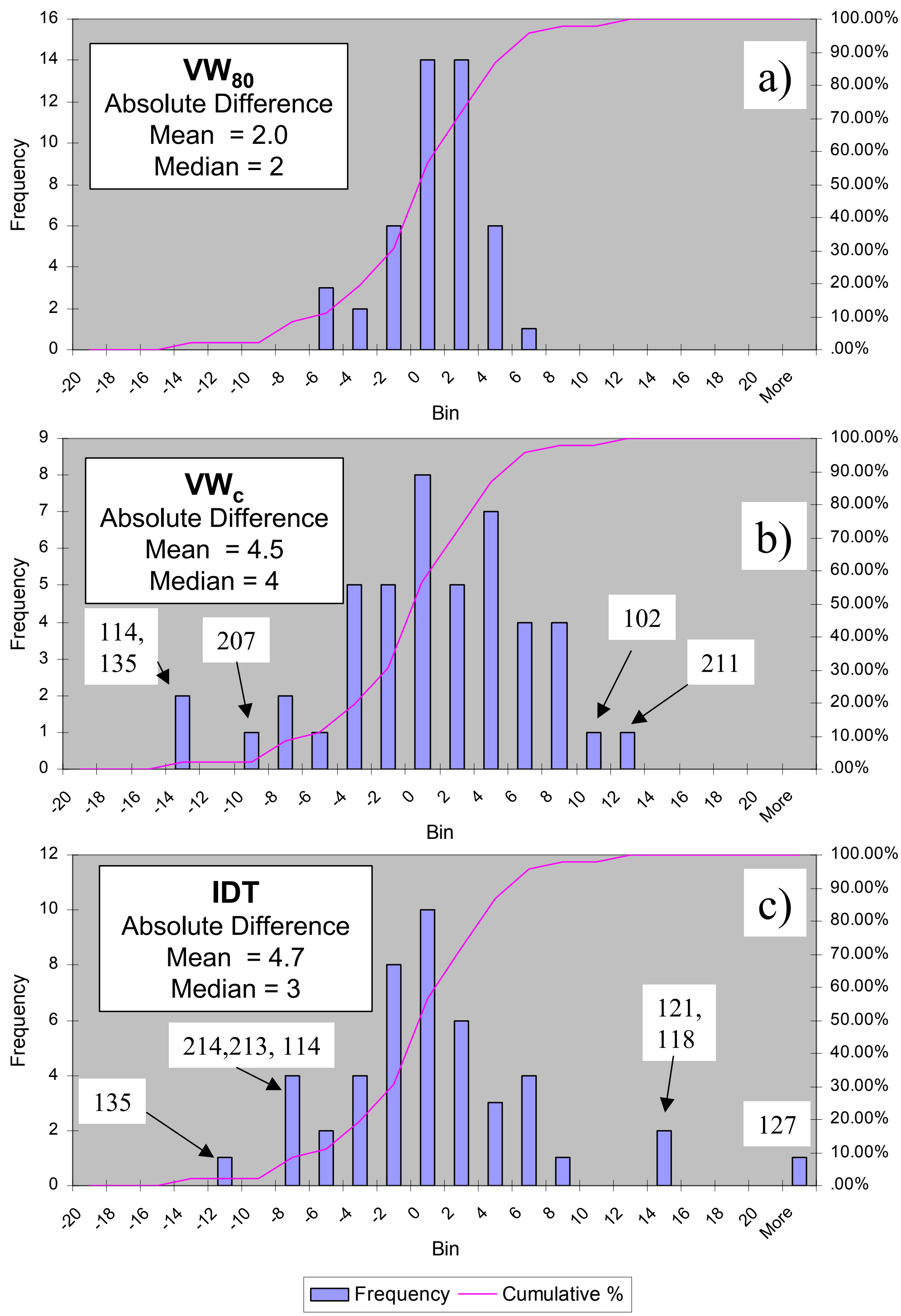


Figure 2-13. Histograms of the Changes in OSF Rankings that Result from Subtracting the Area-Based MOE Rankings from the Nominal MOE Rankings for Summed Concentration Comparisons: a) VW_{80}, b) VW_c, and c) IDT

Table 2-22. Agreements of the Top 10 and Bottom 10 Ranked Models for Summed Concentration and 4 Different Scoring Functions When Compared to the Rankings Based on the Nominal MOE Values

Agreements for Top and Bottom Ten Ranked Models Based on Summed Concentration MOE Values			
	VW_{80}	VWc	IDT
OSF	9 / 9	8 / 8	6 / 8
RWFMS(1,1)	9 / 9	6 / 7	7 / 6
RWFMS(5,0.5)	9 / 9	9 / 6	7 / 6
ABS(FB)	8 / 9	8 / 7	4 / 7

It must be noted that interpolating between sampler locations cannot capture "peaks" or "holes" in the concentration distribution that may lie between samplers. For densely sampled regions this would not be a problem. However, for situations where, for example, complex terrain or a highly urbanized environment lies between perhaps sparse sampler locations, one might expect considerable variations in the concentrations as a function of time and location. Over the long distances associated with ETEX, it is reasonable to expect that the locations of any holes or peaks may shift in time and ultimately be mitigated by dispersive effects.[10] Furthermore, in the next chapter, we start by examining dosage-based MOE values. These dosage-based values consider the summation of thirty 3-hour concentration time periods. The summation process should reduce the likelihood of large unexpected variations (peaks or holes) between sampler locations by smoothing out temporal differences that may be evident at the 3-hour time resolution.

10 Ref. 2-3 (page 3-12) describes evidence for pockets of zero dosage ("holes") at very short range – 50 meters. It is also suggested in Ref. 2-3 that dispersive effects tend to allow the plume to fill in these zero dosage pockets by about 800 meters.

REFERENCE

2-1. Warner S., Platt, N., Heagy, J. F., 2003: *Application of User-Oriented MOE to Transport and Dispersion Model Predictions of the European Tracer Experiment*, IDA Paper P-3829, 86 pp, November 2003. (Available electronically [DTIC STINET ada419433] or on CD via e-mail request to Steve Warner at swarner@ida.org or a mail request to Steve Warner, Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, Virginia 22311-1882); Platt, N., Warner, S., and Heagy, J. F. 2004: Application of User-Oriented MOE to Transport and Dispersion Model Predictions of ETEX. *Proceedings of the Ninth International Conference on Harmonisation Within Atmospheric Dispersion Modelling for Regulatory Purposes*, Garmisch-Partenkirchen, Germany, 1-4 June 2004, pages 120-125; and Warner S., Platt, N., Heagy, J. F., 2004: Application of user-oriented MOE to transport and dispersion model predictions of the European tracer experiment. *Atmospheric Environment*, in press.

2-2. IDL 6.0, 2003: *Interactive Data Language (IDL)* developed by Research Systems Inc., www.rsinc.com.

2-3. Warner, S., Platt, N., and Heagy, 2001: *Application of User-Oriented MOE to HPAC Probabilistic Predictions of* Prairie Grass *Field*, IDA Paper P-3586, 300 pp, May 2001. (Available electronically [DTIC STINET ada391653] or on CD via e-mail request to Steve Warner at swarner@ida.org or a mail request to Steve Warner, Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, Virginia 22311-1882.)

2-4. Guibas, L. J., Knuth, D. E., and Sharir, M., 1992: "Randomized Incremental Construction of Delaunay and Voronoi Diagrams," *Algorithmica* 7: 381-413, 1992. Also, see http://www.gris.uni-tuebingen.de/gris/proj/dt/dteng.html.

CHAPTER 3

RESULTS AND DISCUSSION:
POPULATION-BASED MOE VALUES AND COMPARISONS

3. RESULTS AND DISCUSSION: POPULATION-BASED MOE VALUES AND COMPARISONS

This chapter describes our calculations of population-based MOE values for the 46 sets of predictions of *ETEX*. In order to compute population-based MOE values, two extensions to the efforts of Chapter 2 are required. First, dosage-based MOE values must be created. Next, the underlying European population distribution must be considered.

A. DOSAGES, THRESHOLDS, AND POPULATION DISTRIBUTION

To create "observed" dosages at given locations, one simply sums the concentrations at each sampler location. For example, if a 3-hour average concentration of 0.01 ng m^{-3} were observed for 12 hours (720 min) at a given location, a dosage of 7.2 ng min m^{-3} would be computed for that site. However, periods of time in which sampler data could not be (or were not) collected exist for many of the sampler locations. If there were only a few of these missing points, one could simply remove them (along with the corresponding prediction) from the analysis and only compute dosages for locations that had continuous coverage. For the *ETEX* release, however, there are many locations that have at least some missing time periods. Therefore, one must fill in these values in some manner in order to create a dosage. The spatial interpolation already completed (see Chapter 2) for the IDT area-based MOE values provides a natural way to fill in the temporal holes in the observed concentration data. Since predictions exist (in general) at all time periods, predicted dosages can be created by direct summation of the predicted concentrations. For the few model cases where predictions were missing for some samplers and at some time periods, the corresponding observation was removed from the calculation. This procedure leads to IDT, area-based, dosage MOE values.

Dosage-based MOE values were also created by considering the VW_c area-based MOE values. Observed and predicted dosages can be created by summing concentrations at each of the sampler locations. For missing observations, the 3-hour average concentration was set to 0 ng m^{-3}. To avoid inherent biasing, the corresponding prediction was also set to 0 ng m^{-3}. This procedure leads to VW_c, area-based, dosage MOE values that will be compared with the IDT-based values.

For this analysis, three threshold dosages were examined: 7.2, 72, and 360 ng min m^{-3}. These three values can be related to the 3-hour average concentration thresholds of 0.01, 0.1, and 0.5 ng m^{-3} (Chapter 2) by considering a 12-hour (720 minutes) period in which the cloud might pass over any individual sampler location. That is, 720 min × 0.01 ng m^{-3} = 7.2 ng min m^{-3}; 720 × 0.1 = 72 ng min m^{-3}; and 720 × 0.5 = 360 ng min m^{-3}.

Finally, the dosage-based MOE values can be converted into population-based values by including the underlying non-uniform European population distribution. First, for Voronoi-based computations, let *D(i)* be the dosage at sampler *i* and T_D be a dosage threshold of interest. Then, we identify *OVD(i), FND(i),* and *FPD(i)* as follows:

$$
\begin{aligned}
OVD(i) &= \begin{Bmatrix} 1 & \textit{if observed } D(i) \geq T_D \text{ and } \textit{predicted } D(i) \geq T_D \\ 0 & \textit{otherwise} \end{Bmatrix} \\
FND(i) &= \begin{Bmatrix} 1 & \textit{if observed } D(i) \geq T_D \text{ and } \textit{predicted } D(i) < T_D \\ 0 & \textit{otherwise} \end{Bmatrix}. \\
FPD(i) &= \begin{Bmatrix} 1 & \textit{if observed } D(i) < T_D \text{ and } \textit{predicted } D(i) \geq T_D \\ 0 & \textit{otherwise} \end{Bmatrix}
\end{aligned}
\tag{3-1}
$$

Populations within Voronoi polygons, p_i, are then used as weights in the computations of A_{OV}, A_{FN}, and A_{FP} (in a manner analogous to the use of a_i shown in Chapter 2 for Voronoi area based calculations).

$$
\begin{aligned}
A_{OV} &= \sum_{i=1}^{N} \left(p_i \times OVD(i)\right) \\
A_{FN} &= \sum_{i=1}^{N} \left(p_i \times FND(i)\right) \\
A_{FP} &= \sum_{i=1}^{N} \left(p_i \times FPD(i)\right)
\end{aligned}
\tag{3-2}
$$

where *N* = the number of observation/prediction pairings.

Figure 3-1 illustrates the population distribution that was used. The population distribution shown below is represented by population values at about 2.1 million grid cells; 1501 in the x direction ("east-west") and 1401 in the y direction ("north-south"). This results in a grid cell size of 2 km by 2 km. The overall European population represented here is about 500 million.[1] At this point then, for a given threshold, the MOE

1 The population data were extracted from the Missile Defense Agency's Post Engagement Ground Effects Model version 3.6.0.1 (dated June 2001).

values can be expressed, with the x-axis labeled "one minus the fraction of the population inadvertently exposed" and the y-axis labeled "one minus the fraction of the population unnecessarily warned" – i.e., population-based MOE values.

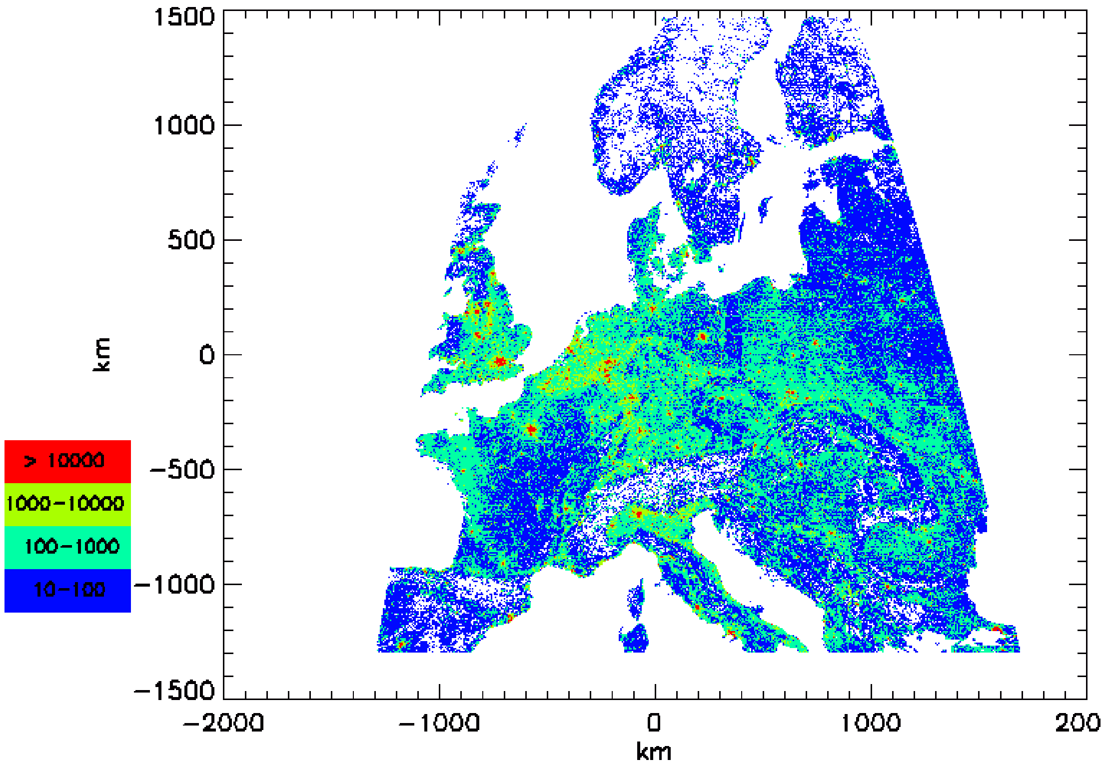


Figure 3-1. Illustration of European Population Distribution

For IDT-based calculations, we define $p_{i,j}$ as the population and $D(i,j)$ as the dosage associated with the 2 km by 2 km grid cell (i,j). Then,

$$OVD(i,j) = \begin{Bmatrix} 1 & \textit{if observed } D(i,j) \geq T_D \text{ and } \textit{predicted } D(i,j) \geq T_D \\ 0 & \textit{otherwise} \end{Bmatrix}$$

$$FND(i,j) = \begin{Bmatrix} 1 & \textit{if observed } D(i,j) \geq T_D \text{ and } \textit{predicted } D(i,j) < T_D \\ 0 & \textit{otherwise} \end{Bmatrix}. \qquad (3\text{-}3)$$

$$FPD(i,j) = \begin{Bmatrix} 1 & \textit{if observed } D(i,j) < T_D \text{ and } \textit{predicted } D(i,j) \geq T_D \\ 0 & \textit{otherwise} \end{Bmatrix}$$

Summing for all grid cells and including the population weights ($p_{i,j}$) leads to values for A_{OV}, A_{FN}, A_{FP} that are based on the European population.

$$A_{OV} = \sum_{i=1}^{N_i}\sum_{j=i}^{N_j}\left(p_{i,j} \times OVD(i,j)\right)$$
$$A_{FN} = \sum_{i=1}^{N_i}\sum_{j=i}^{N_j}\left(p_{i,j} \times FND(i,j)\right). \quad (3\text{-}4)$$
$$A_{FP} = \sum_{i=1}^{N_i}\sum_{j=i}^{N_j}\left(p_{i,j} \times FPD(i,j)\right)$$

B. VORONOI-BASED

Tables 3-1 through 3-9 compare the OSF, RWFMS(1,1), and RWFMS(5,0.5) rankings for three dosage thresholds – 7.2, 72, and 360 ng min m^{-3}. The sixth column in each table presents rankings based on dosage MOE values. The corresponding OSF values are denoted VW_{cDos} because they are based on Voronoi weighting with clipping (as in Chapter 2, VW_c) and they are dosage based (hence the "Dos" subscript). These dosage MOE values can be alternatively thought of as corresponding to population-based values for a uniform population. The last two columns in each table list the rankings based on the actual European population distribution with the associated scoring function values reported in the column labeled VW_{cAP}. For each table, nominal (Ref. 3-1) and VW_c area-based (Chapter 2) rankings are shown for comparisons in the columns 2-3 and 4-5, respectively.

Table 3-10 provides a summary of how model rankings change as a function of MOE type – nominal, VW_c, VW_{cDos}, and VW_{cAP}. The values reported in Table 3-10 correspond to the median difference in red and the mean difference in blue for the absolute rankings for the 46 models for each of the six possible ranking comparisons.

Several results can be obtained from Table 3-10. First, differences in model rankings are greatest when comparing concentration-based and dosage-based MOE values. The middle four columns (3 though 6) compare ranking differences for concentration-based and dosage-based MOE values and result in median ranking differences between 4 and 9.5 with a *median of the medians* of 6 and mean ranking differences between 5.7 and 10.0 with a *median of means* of 7.7. For comparison, median ranking differences for nominal versus VW_c (column 2), *which examine differences due to basing the concentration MOE on areas,* were between 2 and 7 (with a median of the medians of 4) and mean ranking differences between 2.3 and 7.3 (with a median of means of 5.4). Similarly, median ranking differences for VW_{cDos} versus

Table 3-1. Comparisons of Model Rankings Based on OSF: Nominal and VW_c Area-Based for a 0.01 ng m^{-3} Concentration Threshold and for VW_{cDos} and VW_{cAP} for a 7.2 ng min m^{-3} Dosage Threshold

Rank	Model	Nominal	Model	VW_c	Model	VW_{cDos}	Model	VW_{cAP}
1	202	0.358	105	0.308	131	0.174	131	0.154
2	105	0.361	202	0.356	121	0.179	121	0.155
3	208	0.388	131	0.379	208	0.231	205	0.159
4	127	0.389	127	0.395	105	0.239	208	0.162
5	128	0.397	208	0.402	104	0.242	104	0.174
6	210	0.413	210	0.409	205	0.245	101	0.185
7	131	0.420	102	0.415	213	0.250	111	0.187
8	101	0.420	213	0.423	119	0.259	127	0.193
9	205	0.420	204	0.439	118	0.260	105	0.203
10	114	0.424	101	0.444	111	0.289	118	0.204
11	106	0.427	106	0.453	202	0.290	134	0.217
12	110	0.431	110	0.460	210	0.291	110	0.217
13	204	0.439	209	0.461	204	0.303	204	0.222
14	118	0.441	205	0.465	127	0.310	119	0.223
15	209	0.445	128	0.474	203	0.313	202	0.225
16	107	0.451	119	0.476	134	0.316	210	0.233
17	213	0.453	118	0.478	109	0.319	102	0.236
18	113	0.457	111	0.485	101	0.322	213	0.240
19	111	0.463	115	0.510	120	0.322	203	0.248
20	108	0.464	108	0.519	102	0.328	106	0.253
21	116	0.472	121	0.531	110	0.335	132	0.272
22	115	0.485	104	0.551	115	0.338	115	0.284
23	119	0.494	107	0.556	107	0.350	113	0.289
24	121	0.507	134	0.559	106	0.353	109	0.289
25	134	0.508	123	0.563	209	0.355	209	0.298
26	203	0.508	203	0.583	206	0.380	107	0.298
27	123	0.516	116	0.606	108	0.388	116	0.299
28	207	0.519	103	0.607	103	0.389	123	0.315
29	103	0.532	109	0.610	211	0.390	103	0.315
30	104	0.533	201	0.616	123	0.398	120	0.321
31	201	0.542	211	0.617	116	0.403	206	0.324
32	135	0.543	122	0.641	201	0.406	128	0.355
33	102	0.568	120	0.652	132	0.430	108	0.368
34	109	0.569	113	0.654	207	0.445	112	0.380
35	122	0.570	135	0.656	128	0.446	201	0.385
36	112	0.578	207	0.672	135	0.454	130	0.400
37	133	0.579	112	0.686	112	0.475	135	0.402
38	211	0.597	206	0.689	130	0.476	114	0.405
39	120	0.629	133	0.702	122	0.479	207	0.409
40	206	0.648	114	0.726	133	0.493	211	0.417
41	132	0.675	132	0.802	113	0.541	122	0.470
42	214	0.681	214	0.846	114	0.552	133	0.474
43	129	0.883	129	0.902	214	0.561	214	0.582
44	117	0.927	130	0.929	129	0.693	129	0.587
45	130	0.945	212	1.001	117	0.828	117	0.814
46	212	0.974	117	1.038	212	0.905	212	0.828

Table 3-2. Comparisons of Model Rankings Based on RWFMS(1,1): Nominal and VW_c Area-Based for a 0.01 ng m^{-3} Concentration Threshold and for VW_{cDos} and VW_{cAP} for a 7.2 ng min m^{-3} Dosage Threshold

Rank	Model	Nominal	Model	VW_c	Model	VW_{cDos}	Model	VW_{cAP}
1	202	0.597	105	0.642	131	0.781	121	0.804
2	105	0.594	202	0.598	121	0.776	131	0.803
3	208	0.574	131	0.577	208	0.721	205	0.800
4	127	0.568	127	0.565	104	0.720	208	0.797
5	128	0.565	208	0.559	105	0.714	104	0.786
6	210	0.548	210	0.551	205	0.708	111	0.783
7	131	0.546	102	0.547	213	0.699	101	0.772
8	101	0.545	213	0.539	119	0.698	127	0.761
9	205	0.544	101	0.527	118	0.693	118	0.758
10	114	0.541	204	0.527	210	0.662	105	0.755
11	106	0.537	106	0.516	202	0.662	110	0.736
12	110	0.535	110	0.511	111	0.660	119	0.736
13	118	0.530	205	0.510	204	0.652	134	0.734
14	204	0.526	209	0.508	120	0.645	202	0.731
15	209	0.521	128	0.498	127	0.641	204	0.729
16	107	0.517	119	0.497	109	0.639	102	0.719
17	213	0.516	118	0.496	203	0.638	210	0.718
18	113	0.514	111	0.489	101	0.636	203	0.716
19	111	0.507	115	0.469	134	0.635	213	0.710
20	108	0.506	108	0.461	115	0.630	106	0.697
21	116	0.500	121	0.454	110	0.627	113	0.684
22	115	0.489	104	0.437	102	0.623	115	0.680
23	119	0.485	107	0.435	107	0.605	132	0.677
24	121	0.472	134	0.432	106	0.601	123	0.672
25	134	0.471	123	0.423	209	0.600	120	0.670
26	203	0.470	203	0.415	123	0.593	109	0.667
27	123	0.464	116	0.398	206	0.592	107	0.664
28	207	0.461	103	0.393	108	0.592	116	0.655
29	104	0.452	201	0.392	211	0.578	209	0.654
30	103	0.450	109	0.391	103	0.572	206	0.645
31	201	0.445	211	0.380	201	0.560	103	0.636
32	135	0.441	122	0.362	116	0.558	108	0.625
33	109	0.423	135	0.357	132	0.535	128	0.602
34	122	0.422	120	0.356	207	0.523	114	0.588
35	102	0.419	113	0.350	128	0.520	112	0.577
36	112	0.412	207	0.347	135	0.514	201	0.575
37	133	0.409	206	0.333	112	0.498	211	0.571
38	211	0.397	112	0.332	130	0.496	130	0.561
39	120	0.374	133	0.317	122	0.494	135	0.559
40	206	0.363	114	0.299	133	0.482	207	0.553
41	132	0.344	132	0.276	113	0.445	122	0.505
42	214	0.329	214	0.235	114	0.444	133	0.500
43	130	0.188	130	0.206	214	0.430	214	0.416
44	129	0.120	129	0.110	129	0.307	129	0.413
45	117	0.097	117	0.050	117	0.191	117	0.185
46	212	0.072	212	0.043	212	0.100	212	0.172

Table 3-3. Comparisons of Model Rankings Based on RWFMS(5,0.5): Nominal and VW_c Area-Based for a 0.01 ng m^{-3} Concentration Threshold and for VW_{cDos} and VW_{cAP} for a 7.2 ng min m^{-3} Dosage Threshold

Rank	Model	Nominal	Model	VW_c	Model	VW_{cDos}	Model	VW_{cAP}
1	113	0.402	205	0.446	123	0.683	123	0.721
2	101	0.401	101	0.434	104	0.611	114	0.695
3	114	0.396	123	0.396	121	0.587	113	0.643
4	123	0.394	105	0.391	114	0.576	104	0.641
5	135	0.388	110	0.385	131	0.562	205	0.638
6	103	0.384	127	0.377	115	0.559	101	0.605
7	110	0.384	115	0.376	206	0.548	115	0.588
8	205	0.381	202	0.343	205	0.524	131	0.583
9	207	0.355	106	0.342	110	0.510	206	0.580
10	115	0.348	131	0.330	109	0.490	121	0.571
11	116	0.338	204	0.326	101	0.489	110	0.554
12	106	0.334	210	0.295	210	0.488	208	0.552
13	112	0.327	104	0.288	204	0.482	210	0.521
14	127	0.325	213	0.287	113	0.471	127	0.514
15	105	0.322	113	0.286	213	0.471	109	0.513
16	202	0.316	209	0.283	208	0.452	134	0.501
17	203	0.314	102	0.274	201	0.447	204	0.496
18	204	0.289	208	0.271	127	0.430	116	0.481
19	109	0.281	111	0.267	105	0.425	111	0.473
20	104	0.277	109	0.266	103	0.415	213	0.472
21	107	0.276	103	0.265	134	0.410	105	0.468
22	210	0.274	121	0.251	111	0.406	118	0.456
23	214	0.270	112	0.243	209	0.404	209	0.454
24	209	0.269	135	0.241	118	0.403	202	0.440
25	208	0.268	203	0.238	203	0.396	106	0.437
26	128	0.263	114	0.234	119	0.390	132	0.433
27	201	0.254	116	0.229	202	0.380	119	0.432
28	206	0.249	206	0.229	207	0.378	102	0.424
29	121	0.240	128	0.227	106	0.377	103	0.409
30	131	0.240	207	0.225	102	0.374	201	0.407
31	108	0.239	107	0.220	116	0.369	203	0.390
32	111	0.238	201	0.220	112	0.366	135	0.382
33	213	0.222	119	0.205	132	0.361	130	0.381
34	118	0.217	118	0.201	135	0.337	207	0.374
35	134	0.192	134	0.187	107	0.319	112	0.371
36	119	0.179	108	0.178	130	0.313	107	0.335
37	122	0.167	214	0.150	120	0.305	128	0.307
38	102	0.149	211	0.130	128	0.279	120	0.297
39	211	0.140	122	0.122	214	0.252	108	0.255
40	133	0.138	120	0.122	211	0.251	214	0.222
41	120	0.133	132	0.121	108	0.242	211	0.219
42	132	0.123	133	0.103	122	0.207	133	0.202
43	130	0.061	130	0.081	133	0.204	122	0.195
44	129	0.027	129	0.025	129	0.082	129	0.124
45	117	0.022	117	0.011	117	0.048	117	0.044
46	212	0.016	212	0.009	212	0.022	212	0.040

Table 3-4. Comparisons of Model Rankings Based on OSF: Nominal and VW_c Area-Based for a 0.1 ng m^{-3} Concentration Threshold and for VW_{cDos} and VW_{cAP} for a 72 ng min m^{-3} Dosage Threshold

Rank	Model	Nominal	Model	VW_c	Model	VW_{cDos}	Model	VW_{cAP}
1	208	0.381	107	0.502	105	0.159	105	0.144
2	128	0.411	101	0.502	208	0.171	127	0.158
3	202	0.419	127	0.515	119	0.172	119	0.170
4	101	0.424	202	0.515	127	0.199	208	0.171
5	127	0.440	105	0.519	110	0.231	104	0.209
6	107	0.446	131	0.520	202	0.251	110	0.214
7	105	0.451	115	0.548	210	0.252	101	0.223
8	131	0.462	208	0.563	205	0.263	205	0.223
9	118	0.476	128	0.567	209	0.278	102	0.234
10	115	0.481	213	0.570	111	0.287	202	0.238
11	205	0.488	134	0.576	121	0.291	210	0.268
12	134	0.492	204	0.580	102	0.300	209	0.274
13	106	0.494	111	0.582	101	0.300	121	0.289
14	210	0.495	209	0.582	104	0.304	132	0.298
15	114	0.499	102	0.584	128	0.349	213	0.312
16	111	0.505	118	0.598	134	0.355	111	0.314
17	209	0.509	106	0.599	115	0.356	115	0.318
18	204	0.513	211	0.606	213	0.358	103	0.326
19	213	0.522	203	0.610	211	0.385	128	0.332
20	110	0.526	210	0.630	103	0.387	134	0.336
21	133	0.526	123	0.644	131	0.392	131	0.340
22	119	0.547	104	0.655	204	0.418	204	0.346
23	113	0.560	121	0.666	123	0.419	123	0.351
24	207	0.562	108	0.673	106	0.426	106	0.361
25	123	0.565	110	0.676	107	0.434	206	0.363
26	102	0.572	119	0.698	207	0.444	112	0.381
27	201	0.594	120	0.699	122	0.444	113	0.382
28	203	0.606	207	0.701	201	0.473	211	0.392
29	211	0.612	122	0.716	206	0.479	207	0.406
30	121	0.637	205	0.722	112	0.484	107	0.422
31	108	0.638	114	0.745	118	0.507	114	0.424
32	104	0.647	113	0.746	109	0.528	201	0.449
33	103	0.652	201	0.793	203	0.552	109	0.465
34	122	0.653	109	0.811	120	0.554	118	0.468
35	120	0.671	133	0.835	108	0.577	203	0.481
36	135	0.687	103	0.836	132	0.585	120	0.483
37	116	0.694	206	0.846	113	0.597	122	0.488
38	109	0.695	112	0.858	130	0.612	130	0.504
39	112	0.741	214	0.891	114	0.612	129	0.561
40	206	0.778	129	0.899	129	0.642	108	0.566
41	132	0.803	132	0.919	116	0.672	133	0.586
42	214	0.817	116	0.949	214	0.730	116	0.614
43	129	0.822	117	1.045	133	0.762	135	0.616
44	117	0.927	212	1.100	135	0.808	214	0.730
45	212	1.071	135	1.110	117	0.879	117	0.809
46	130	1.092	130	1.153	212	0.903	212	0.873

Table 3-5. Comparisons of Model Rankings Based on RWFMS(1,1): Nominal and VW_c Area-Based for a 0.1 ng m^{-3} Concentration Threshold and for VW_{cDos} and VW_{cAP} for a 72 ng min m^{-3} Dosage Threshold

Rank	Model	Nominal	Model	VW_c	Model	VW_{cDos}	Model	VW_{cAP}
1	208	0.577	107	0.476	105	0.800	105	0.829
2	128	0.551	101	0.476	208	0.791	119	0.810
3	202	0.545	127	0.466	119	0.791	127	0.799
4	101	0.544	202	0.464	127	0.755	208	0.787
5	127	0.526	131	0.462	110	0.727	104	0.750
6	107	0.521	105	0.461	210	0.703	110	0.745
7	105	0.517	115	0.438	202	0.699	205	0.736
8	131	0.508	208	0.430	205	0.697	101	0.733
9	118	0.497	128	0.427	111	0.686	102	0.716
10	115	0.493	213	0.426	209	0.679	202	0.712
11	205	0.487	134	0.421	101	0.662	210	0.684
12	134	0.484	111	0.417	121	0.662	209	0.679
13	106	0.482	204	0.415	104	0.661	121	0.672
14	210	0.481	209	0.414	102	0.655	111	0.663
15	114	0.478	102	0.412	115	0.607	132	0.657
16	111	0.473	211	0.399	128	0.605	213	0.641
17	209	0.470	203	0.397	213	0.602	115	0.639
18	204	0.467	106	0.396	134	0.599	103	0.625
19	213	0.460	118	0.395	211	0.575	128	0.625
20	110	0.456	210	0.383	103	0.572	131	0.619
21	133	0.455	121	0.358	131	0.572	134	0.617
22	119	0.439	104	0.358	123	0.555	204	0.614
23	113	0.428	123	0.353	204	0.544	123	0.614
24	207	0.426	108	0.345	106	0.538	206	0.600
25	102	0.423	119	0.339	107	0.533	106	0.594
26	123	0.421	110	0.338	122	0.525	211	0.582
27	201	0.403	207	0.334	207	0.523	113	0.577
28	203	0.399	120	0.333	201	0.499	112	0.576
29	211	0.396	122	0.318	206	0.495	107	0.559
30	121	0.379	205	0.309	112	0.490	207	0.554
31	108	0.378	114	0.305	118	0.472	114	0.541
32	122	0.368	113	0.294	109	0.456	201	0.518
33	104	0.365	201	0.280	203	0.434	118	0.516
34	120	0.354	109	0.271	120	0.432	109	0.505
35	103	0.351	103	0.248	132	0.413	120	0.499
36	135	0.345	133	0.246	108	0.411	203	0.498
37	116	0.341	112	0.242	113	0.397	122	0.489
38	109	0.336	206	0.232	114	0.393	130	0.473
39	112	0.306	214	0.227	130	0.391	129	0.429
40	132	0.270	132	0.211	129	0.354	108	0.422
41	206	0.269	116	0.195	116	0.350	133	0.407
42	214	0.263	129	0.157	214	0.318	135	0.392
43	129	0.190	135	0.118	133	0.284	116	0.389
44	130	0.129	130	0.099	135	0.269	214	0.316
45	117	0.125	117	0.067	117	0.161	117	0.191
46	212	0.060	212	0.036	212	0.097	212	0.127

Table 3-6. Comparisons of Model Rankings Based on RWFMS(5,0.5): Nominal and VW_c Area-Based for a 0.1 ng m^{-3} Concentration Threshold and for VW_{cDos} and VW_{cAP} for a 72 ng min m^{-3} Dosage Threshold

Rank	Model	Nominal	Model	VW_c	Model	VW_{cDos}	Model	VW_{cAP}
1	101	0.438	101	0.395	110	0.591	104	0.610
2	110	0.409	105	0.350	205	0.579	110	0.609
3	123	0.385	202	0.340	104	0.568	205	0.604
4	205	0.381	123	0.330	127	0.562	127	0.581
5	208	0.380	115	0.311	105	0.553	101	0.576
6	202	0.377	106	0.298	101	0.550	105	0.550
7	115	0.357	127	0.290	210	0.540	208	0.528
8	106	0.357	131	0.290	209	0.533	123	0.515
9	105	0.350	102	0.271	208	0.518	210	0.497
10	128	0.344	204	0.265	119	0.513	119	0.496
11	127	0.339	209	0.264	123	0.512	209	0.496
12	207	0.323	110	0.251	115	0.501	206	0.488
13	114	0.309	107	0.246	102	0.487	132	0.488
14	113	0.305	104	0.245	213	0.463	115	0.482
15	209	0.301	208	0.241	202	0.463	102	0.481
16	107	0.300	134	0.230	131	0.446	131	0.480
17	103	0.294	128	0.226	206	0.399	202	0.455
18	210	0.282	205	0.215	103	0.385	213	0.446
19	201	0.275	213	0.213	121	0.373	113	0.393
20	204	0.271	210	0.203	134	0.364	114	0.379
21	131	0.265	113	0.202	111	0.336	103	0.363
22	134	0.246	111	0.200	207	0.333	134	0.347
23	104	0.238	121	0.196	128	0.329	121	0.346
24	213	0.227	203	0.186	112	0.326	207	0.328
25	111	0.221	207	0.181	204	0.321	112	0.325
26	203	0.217	211	0.179	106	0.321	130	0.320
27	118	0.207	114	0.172	113	0.314	128	0.316
28	109	0.198	206	0.155	201	0.293	106	0.315
29	102	0.196	119	0.150	211	0.284	111	0.307
30	206	0.193	103	0.142	130	0.254	204	0.298
31	211	0.188	201	0.141	132	0.238	109	0.258
32	121	0.185	118	0.135	107	0.238	201	0.253
33	133	0.183	120	0.125	114	0.231	211	0.242
34	135	0.182	109	0.125	122	0.222	107	0.218
35	112	0.180	108	0.123	109	0.209	118	0.187
36	119	0.176	112	0.120	118	0.176	203	0.183
37	122	0.166	122	0.114	203	0.167	122	0.183
38	108	0.166	214	0.103	120	0.165	120	0.180
39	116	0.154	133	0.083	108	0.137	135	0.169
40	120	0.145	132	0.081	116	0.132	133	0.148
41	214	0.140	116	0.074	214	0.131	116	0.147
42	132	0.099	130	0.046	135	0.101	129	0.136
43	130	0.051	135	0.040	129	0.101	108	0.136
44	129	0.047	129	0.039	133	0.094	214	0.124
45	117	0.030	117	0.015	117	0.040	117	0.046
46	212	0.014	212	0.008	212	0.021	212	0.028

Table 3-7. Comparisons of Model Rankings Based on OSF: Nominal and VW_c Area-Based for a 0.5 ng m^{-3} Concentration Threshold and for VW_{cDos} and VW_{cAP} for a 360 ng min m^{-3} Dosage Threshold

Rank	Model	Nominal	Model	VW_c	Model	VW_{cDos}	Model	VW_{cAP}
1	127	0.600	127	0.633	127	0.447	127	0.263
2	107	0.632	118	0.709	113	0.450	208	0.364
3	134	0.635	205	0.811	128	0.473	105	0.369
4	118	0.646	113	0.822	209	0.485	128	0.370
5	128	0.658	209	0.823	208	0.530	107	0.393
6	208	0.676	128	0.834	108	0.534	213	0.394
7	111	0.676	208	0.834	107	0.538	209	0.399
8	133	0.682	107	0.859	134	0.538	204	0.406
9	131	0.687	210	0.868	111	0.582	122	0.420
10	205	0.704	134	0.868	105	0.586	120	0.423
11	209	0.709	110	0.870	110	0.593	113	0.425
12	105	0.714	131	0.873	210	0.594	210	0.437
13	119	0.724	111	0.874	213	0.598	111	0.446
14	110	0.735	105	0.894	203	0.614	134	0.450
15	101	0.746	101	0.900	101	0.614	109	0.455
16	210	0.750	213	0.908	205	0.616	103	0.456
17	202	0.750	123	0.909	204	0.619	101	0.474
18	113	0.752	202	0.922	104	0.646	203	0.475
19	106	0.756	115	0.957	109	0.647	104	0.479
20	213	0.766	133	0.958	202	0.650	115	0.479
21	123	0.804	106	0.961	121	0.653	202	0.479
22	207	0.807	104	0.968	133	0.658	112	0.484
23	121	0.822	211	0.969	131	0.661	106	0.488
24	204	0.832	201	0.969	122	0.683	131	0.498
25	203	0.841	103	0.972	106	0.694	110	0.505
26	201	0.846	122	0.974	103	0.734	205	0.506
27	102	0.870	204	0.989	118	0.741	211	0.515
28	104	0.874	119	0.998	115	0.748	118	0.521
29	115	0.876	121	1.000	123	0.752	108	0.526
30	116	0.879	108	1.011	102	0.782	121	0.527
31	211	0.920	207	1.021	112	0.788	201	0.547
32	120	0.922	120	1.024	211	0.809	133	0.547
33	122	0.928	203	1.025	120	0.843	119	0.568
34	132	0.948	102	1.046	119	0.849	206	0.568
35	103	0.954	112	1.054	201	0.853	102	0.624
36	108	0.977	132	1.070	206	0.854	207	0.634
37	206	0.993	206	1.079	207	0.903	123	0.643
38	112	0.995	116	1.102	117	0.920	117	0.674
39	114	1.002	114	1.103	129	0.922	129	0.681
40	129	1.025	109	1.160	214	0.993	116	0.724
41	214	1.028	129	1.175	130	1.010	214	0.753
42	135	1.052	212	1.191	132	1.011	132	0.805
43	117	1.122	117	1.200	114	1.015	135	0.859
44	109	1.156	214	1.205	212	1.067	130	0.883
45	212	1.173	135	1.207	116	1.123	114	0.906
46	130	1.210	130	1.233	135	1.128	212	1.065

Table 3-8. Comparisons of Model Rankings Based on RWFMS(1,1): Nominal and VW_c Area-Based for a 0.5 ng m^{-3} Concentration Threshold and for VW_{cDos} and VW_{cAP} for a 360 ng min m^{-3} Dosage Threshold

Rank	Model	Nominal	Model	VW_c	Model	VW_{cDos}	Model	VW_{cAP}
1	127	0.401	127	0.380	127	0.529	127	0.694
2	107	0.381	118	0.321	113	0.517	105	0.621
3	134	0.371	113	0.260	128	0.510	128	0.607
4	118	0.368	209	0.242	209	0.494	208	0.598
5	111	0.352	205	0.234	208	0.452	107	0.567
6	133	0.349	210	0.232	108	0.448	209	0.565
7	128	0.344	111	0.228	134	0.448	213	0.565
8	131	0.341	107	0.223	107	0.447	204	0.555
9	208	0.338	208	0.216	111	0.414	122	0.543
10	209	0.325	101	0.216	210	0.407	120	0.540
11	119	0.323	131	0.203	105	0.407	113	0.538
12	205	0.319	128	0.194	110	0.401	210	0.528
13	101	0.309	213	0.194	213	0.399	134	0.525
14	210	0.305	115	0.191	203	0.393	111	0.520
15	113	0.300	134	0.186	205	0.382	103	0.517
16	105	0.299	211	0.177	101	0.380	109	0.514
17	213	0.290	123	0.171	204	0.377	202	0.501
18	110	0.278	133	0.165	109	0.372	104	0.499
19	202	0.272	103	0.163	133	0.356	101	0.498
20	106	0.267	121	0.163	104	0.352	203	0.497
21	121	0.265	108	0.161	121	0.350	106	0.496
22	207	0.264	110	0.159	122	0.348	115	0.494
23	123	0.256	122	0.159	131	0.344	205	0.491
24	204	0.255	120	0.153	202	0.344	112	0.490
25	203	0.253	204	0.152	115	0.307	110	0.488
26	115	0.235	203	0.150	103	0.304	131	0.480
27	116	0.232	201	0.150	106	0.303	211	0.466
28	201	0.231	112	0.144	112	0.283	118	0.460
29	102	0.230	119	0.144	211	0.272	121	0.457
30	104	0.212	104	0.139	118	0.264	108	0.456
31	120	0.207	105	0.137	123	0.253	133	0.441
32	211	0.201	132	0.136	120	0.249	201	0.439
33	122	0.198	207	0.130	102	0.242	119	0.422
34	132	0.196	114	0.124	119	0.239	206	0.420
35	108	0.183	202	0.115	201	0.234	207	0.378
36	103	0.180	116	0.111	206	0.226	102	0.371
37	112	0.170	109	0.094	207	0.218	123	0.353
38	114	0.168	206	0.091	129	0.208	129	0.349
39	214	0.157	102	0.086	117	0.202	117	0.329
40	129	0.155	106	0.086	214	0.173	116	0.323
41	206	0.148	129	0.084	132	0.165	214	0.305
42	135	0.143	117	0.080	114	0.164	132	0.270
43	117	0.103	135	0.070	130	0.115	135	0.244
44	109	0.099	214	0.065	135	0.112	114	0.213
45	130	0.072	212	0.044	116	0.111	130	0.187
46	212	0.059	130	0.032	212	0.070	212	0.077

Table 3-9. Comparisons of Model Rankings Based on RWFMS(5,0.5): Nominal and VW_c Area-Based for a 0.5 ng m^{-3} Concentration Threshold and for VW_{cDos} and VW_{cAP} for a 360 ng min m^{-3} Dosage Threshold

Rank	Model	Nominal	Model	VW_c	Model	VW_{cDos}	Model	VW_{cAP}
1	128	0.301	128	0.213	128	0.536	105	0.702
2	110	0.273	127	0.207	205	0.531	205	0.640
3	105	0.270	205	0.198	110	0.512	110	0.602
4	134	0.261	110	0.192	105	0.503	128	0.594
5	208	0.256	208	0.189	127	0.488	127	0.550
6	127	0.246	134	0.177	134	0.473	106	0.519
7	202	0.245	105	0.175	209	0.459	202	0.496
8	106	0.242	209	0.170	202	0.414	134	0.466
9	205	0.235	131	0.158	208	0.398	208	0.465
10	107	0.211	202	0.152	106	0.393	104	0.459
11	131	0.200	107	0.152	101	0.329	209	0.440
12	209	0.198	123	0.148	104	0.326	103	0.435
13	133	0.177	113	0.140	204	0.319	131	0.431
14	123	0.175	213	0.133	123	0.300	204	0.366
15	213	0.169	210	0.128	107	0.295	122	0.357
16	113	0.169	111	0.126	121	0.293	123	0.351
17	210	0.159	104	0.121	131	0.289	213	0.348
18	201	0.156	106	0.120	113	0.285	101	0.344
19	207	0.155	133	0.116	213	0.278	206	0.343
20	111	0.153	201	0.116	111	0.259	102	0.334
21	104	0.147	101	0.115	133	0.241	210	0.322
22	101	0.146	118	0.111	102	0.236	201	0.320
23	118	0.143	122	0.109	203	0.220	109	0.314
24	119	0.142	103	0.107	103	0.200	119	0.312
25	204	0.132	204	0.103	210	0.180	113	0.311
26	102	0.128	119	0.102	122	0.170	115	0.308
27	211	0.115	211	0.098	108	0.168	120	0.291
28	121	0.114	207	0.095	109	0.164	107	0.273
29	122	0.111	115	0.090	206	0.154	121	0.268
30	103	0.107	102	0.089	115	0.150	111	0.261
31	120	0.103	121	0.087	201	0.145	133	0.258
32	206	0.103	203	0.081	119	0.142	211	0.243
33	115	0.102	120	0.080	120	0.129	203	0.231
34	203	0.102	206	0.078	211	0.122	112	0.222
35	116	0.090	116	0.064	112	0.115	207	0.221
36	132	0.073	132	0.063	207	0.108	129	0.178
37	108	0.071	108	0.055	130	0.104	118	0.168
38	114	0.060	114	0.051	129	0.100	108	0.162
39	112	0.060	112	0.051	214	0.080	130	0.161
40	214	0.057	129	0.045	132	0.072	116	0.146
41	129	0.052	214	0.041	118	0.072	214	0.141
42	135	0.049	135	0.039	117	0.067	135	0.102
43	130	0.037	130	0.032	114	0.065	117	0.101
44	109	0.034	109	0.030	116	0.053	132	0.101
45	117	0.030	117	0.027	135	0.046	114	0.074
46	212	0.014	212	0.010	212	0.016	212	0.018

Table 3-10. Median / Mean Absolute Ranking Difference Between Various MOE Value (Nominal, VW_c, VW_{cDos}, and VW_{cAP}) Rankings for 3 Thresholds and for 3 Scoring Functions

Threshold (ng min m^{-3}), Scoring Function	Nominal - VW_c	Nominal - VW_{cDos}	Nominal - VW_{cAP}	VW_c - VW_{cDos}	VW_c - VW_{cAP}	VW_{cDos} - VW_{cAP}
7.2, OSF	2 / 4.3	7.5 / 8.9	6 / 7.9	6.5 / 6.5	5 / 6.7	3 / 4.0
7.2, RWFMS(1,1)	2 / 4.1	7 / 8.7	7 / 8.1	5 / 6.4	6 / 7.0	3.5 / 4.4
7.2, RWFMS(5,0.5)	7 / 7.3	9.5 / 10.0	8.5 / 9.5	5 / 7.0	6 / 7.2	2 / 3.4
72, OSF	4.5 / 5.5	7 / 8.3	6.5 / 9.0	6.5 / 7.7	7.5 / 8.9	2 / 3.3
72, RWFMS(1,1)	4 / 5.4	6 / 8.3	6 / 9.1	6 / 7.4	7 / 8.6	2 / 3.2
72, RWFMS(5,0.5)	5 / 5.7	7.5 / 8.0	5.5 / 8.0	5 / 7.0	6.5 / 7.7	2 / 3.0
360, OSF	3.5 / 4.6	5 / 7.3	7 / 9.4	4 / 5.9	4.5 / 7.5	3.5 / 5.2
360, RWFMS(1,1)	5.5 / 6.7	5 / 7.7	7 / 9.2	6.5 / 7.0	6.5 / 8.7	3 / 5.0
360, RWFMS(5,0.5)	2 / 2.3	4 / 5.7	4 / 6.0	4 / 5.4	4 / 6.0	3.5 / 4.5

VW_{cAP} (column 7), *which examine differences due to basing the dosage MOE on actual European population distributions,* were between 2 and 3.5 (with a median of the medians of 3) and mean ranking differences between 3.0 and 5.2 (with a median of means of 4.0).

A few models improve their relative rankings greatly when assessed based on dosage MOE values instead of concentration-based values. For example, for OSF rankings and the three comparative concentration/dosage thresholds (0.01 ng m^{-3} / 7.2 ng min m^{-3}, 0.1 ng m^{-3} / 72 ng min m-3, 0.5 ng m^{-3} / 360 ng min m^{-3}), model 121 (SCIPUFF) moves up 19 (from 21 to 2, pink and green columns in Table 3-1), 12 (from 23 to 11, Table 3-4), and 8 (from 29 to 21, Table 3-7) positions, respectively. Examination of 3-hour concentration and cumulative dosage plots for the observations and predictions, suggest that some models do not match the 3-hour timing (e.g., time of arrival and dwell) as well as others. Therefore, while dosages may be well predicted, 3-hour average concentrations that require both the location *and* time to be matched may be predicted worse (relative to the other models). This certainly appears to be the case for the SCIPUFF predictions (model 121).

Figure 3-2 shows contours associated with the SCIPUFF 3-hour average concentration predictions and the corresponding observations for the period of time starting one day after the release.[2] The figure indicates that the SCIPUFF predictions seemed to "run ahead" of the observations for the period of time after about 42 hours. Such a mismatch in timing would be expected to degrade 3-hour average concentration MOE values (and hence rankings) for SCIPUFF. However, summing these concentrations over all time periods to create dosage MOE values would result in improved *relative* performance given that other models did not have such timing mismatches.

Table 3-10 and inspection of Tables 3-1 through 3-9 indicate that ranking changes are quite similar for the OSF and RWFMS(1,1) scoring functions (as also described in Chapter 2 for area-based MOE value rankings). Finally, no strong trend in the magnitude of the absolute ranking differences between different MOE computation techniques (nominal, area-, dosage-, or population-based) can be assigned to increases in the threshold.

C. INTERPOLATION-BASED: IDT

Tables 3-11 through 3-19 compare the OSF, RWFMS(1,1), and RWFMS(5,0.5) rankings for three dosage thresholds – 7.2, 72, and 360 ng min m^{-3}. The sixth column in each table presents rankings based on dosage MOE values. These MOE values are denoted IDT_{Dos} because they are based on interpolation (after Delaunay triangulation) (as in Chapter 2, IDT) and they are dosage based (hence the "Dos" subscript). These dosage MOE values can be alternatively thought of as corresponding to population-based values for a uniform population. The last two columns in each table list the model rankings based on the actual European population distribution with the associated scoring function value reported in the column labeled IDT_{AP}. For each table, nominal (Ref. 3-1) and the IDT area-based (Chapter 2) rankings are shown for comparisons in the columns 2-3 and 4-5, respectively.

Table 3-20 provides a summary of how model rankings change as a function of MOE type – nominal, IDT, IDT_{Dos}, and IDT_{AP}. The values reported in Table 3-20 correspond to the median difference in red and the mean difference in blue for the

[2] Both contours were created by using Delaunay triangulation, followed by interpolation – the "IDT" method.

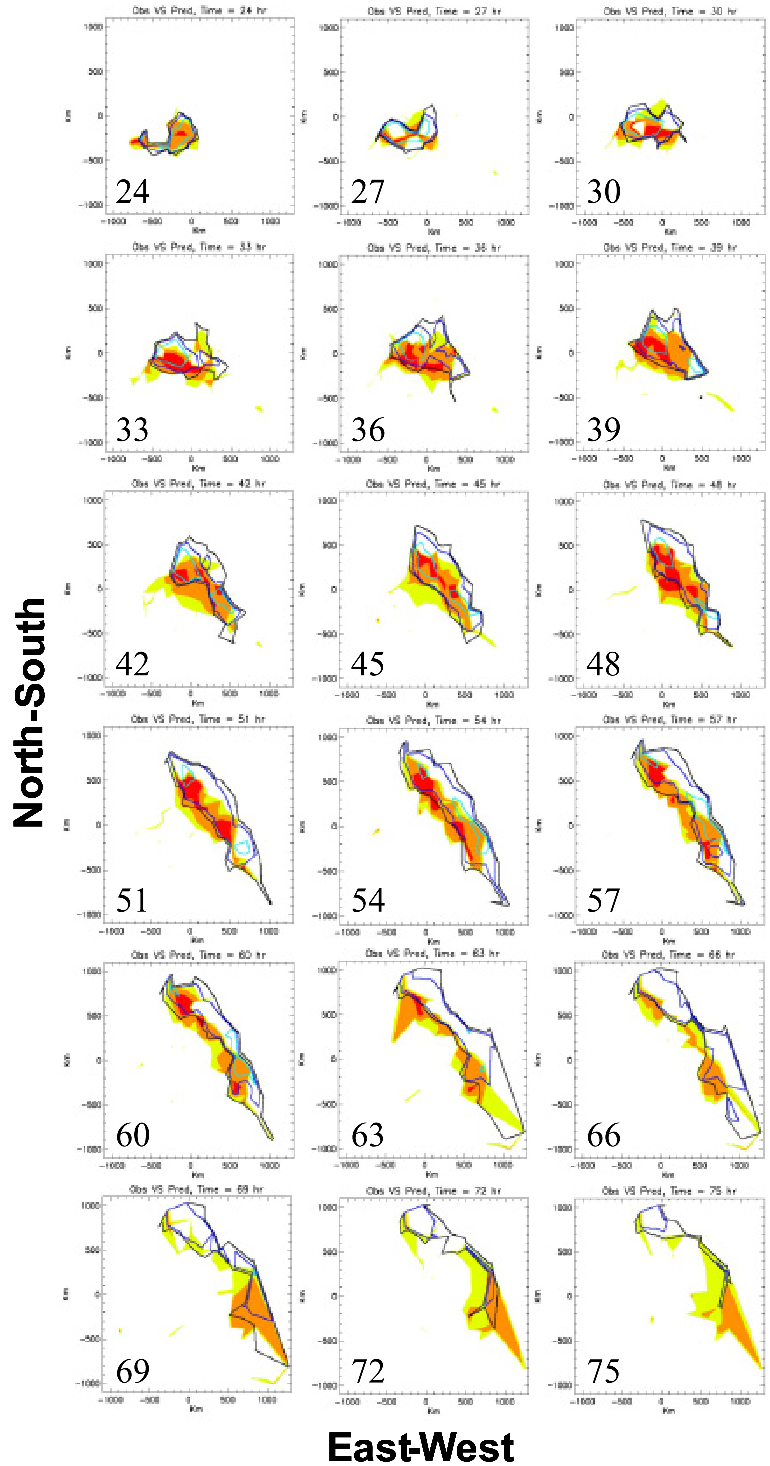


Figure 3-2. Contours (Based on IDT) for 3-hour Average Concentration Observations and SCIPUFF (Model 121) Predictions for the Time Periods Between 36 and 75 Hours After the Release. The solid lines correspond to contours for the SCIPUFF predictions (black = 0.5 ng m^{-3}, dark blue = 0.1 ng m^{-3}, and lighter blue = 0.01 ng m^{-3}) and the shaded regions correspond to "observed" (after IDT procedure) areas above the 3-hour concentration thresholds (red = 0.5 ng m^{-3}, orange = 0.1 ng m^{-3}, and yellow = 0.01 ng m^{-3}). Numbers on individual plots correspond to the last hour of the given 3-hour time period.

Table 3-11. Comparisons of Model Rankings Based on OSF: Nominal and IDT Area-Based for a 0.01 ng m^{-3} Concentration Threshold and for IDT_{Dos} and IDT_{AP} for a 7.2 ng min m^{-3} Dosage Threshold

Rank	Model	Nominal	Model	IDT	Model	IDT_{Dos}	Model	IDT_{AP}
1	202	0.358	105	0.360	208	0.136	208	0.060
2	105	0.361	202	0.415	105	0.160	127	0.078
3	208	0.388	208	0.449	202	0.183	105	0.089
4	127	0.389	127	0.461	134	0.194	121	0.091
5	128	0.397	106	0.473	131	0.197	106	0.098
6	210	0.413	114	0.491	111	0.207	131	0.107
7	131	0.420	113	0.495	121	0.222	110	0.112
8	101	0.420	101	0.509	127	0.223	203	0.116
9	205	0.420	210	0.511	119	0.227	111	0.116
10	114	0.424	128	0.513	210	0.234	134	0.117
11	106	0.427	131	0.516	118	0.237	204	0.120
12	110	0.431	204	0.517	213	0.246	202	0.127
13	204	0.439	205	0.525	128	0.248	104	0.134
14	118	0.441	110	0.538	106	0.251	205	0.141
15	209	0.445	213	0.545	203	0.270	101	0.149
16	107	0.451	209	0.552	116	0.271	128	0.150
17	213	0.453	115	0.568	209	0.272	113	0.151
18	113	0.457	118	0.573	107	0.279	210	0.152
19	111	0.463	111	0.588	204	0.280	116	0.153
20	108	0.464	102	0.592	104	0.284	107	0.157
21	116	0.472	119	0.593	205	0.288	119	0.159
22	115	0.485	107	0.603	110	0.306	112	0.160
23	119	0.494	116	0.606	113	0.309	103	0.170
24	121	0.507	123	0.613	109	0.315	207	0.173
25	134	0.508	207	0.614	115	0.322	115	0.174
26	203	0.508	108	0.617	108	0.322	109	0.179
27	123	0.516	203	0.641	135	0.322	135	0.179
28	207	0.519	103	0.644	207	0.323	213	0.182
29	103	0.532	135	0.646	102	0.333	209	0.187
30	104	0.533	109	0.664	206	0.337	206	0.192
31	201	0.542	134	0.669	114	0.341	123	0.200
32	135	0.543	112	0.686	103	0.342	118	0.204
33	102	0.568	121	0.694	123	0.346	114	0.211
34	109	0.569	133	0.698	101	0.346	108	0.233
35	122	0.570	201	0.705	112	0.349	102	0.244
36	112	0.578	104	0.715	201	0.385	201	0.248
37	133	0.579	122	0.745	211	0.391	132	0.259
38	211	0.597	211	0.753	120	0.440	120	0.329
39	120	0.629	206	0.764	122	0.442	211	0.331
40	206	0.648	120	0.802	214	0.454	122	0.341
41	132	0.675	132	0.851	133	0.467	214	0.351
42	214	0.681	214	0.889	132	0.482	133	0.366
43	129	0.883	129	0.946	130	0.529	130	0.452
44	117	0.927	117	1.019	129	0.740	129	0.667
45	130	0.945	212	1.037	117	0.879	117	0.808
46	212	0.974	130	1.101	212	0.936	212	0.895

Table 3-12. Comparisons of Model Rankings Based on RWFMS(1,1): Nominal and IDT Area-Based for a 0.01 ng m^{-3} Concentration Threshold and for IDT_{Dos} and IDT_{AP} for a 7.2 ng min m^{-3} Dosage Threshold

Rank	Model	Nominal	Model	IDT	Model	IDT_{Dos}	Model	IDT_{AP}
1	202	0.597	105	0.594	208	0.825	208	0.919
2	105	0.594	202	0.546	105	0.801	127	0.904
3	208	0.574	208	0.520	202	0.773	105	0.882
4	127	0.568	127	0.508	134	0.759	121	0.881
5	128	0.565	106	0.499	131	0.759	110	0.877
6	210	0.548	114	0.486	127	0.748	106	0.872
7	131	0.546	113	0.482	111	0.747	131	0.860
8	101	0.545	101	0.469	121	0.741	104	0.858
9	205	0.544	210	0.469	119	0.731	111	0.857
10	114	0.541	128	0.466	118	0.725	134	0.853
11	106	0.537	131	0.464	210	0.721	205	0.851
12	110	0.535	204	0.464	106	0.715	203	0.851
13	118	0.530	205	0.457	205	0.707	204	0.846
14	204	0.526	110	0.445	128	0.705	101	0.841
15	209	0.521	213	0.442	213	0.704	113	0.841
16	107	0.517	209	0.438	104	0.703	202	0.840
17	213	0.516	115	0.422	204	0.687	112	0.825
18	113	0.514	118	0.414	209	0.686	128	0.824
19	111	0.507	111	0.413	203	0.684	119	0.815
20	108	0.506	102	0.408	110	0.682	107	0.813
21	116	0.500	107	0.402	113	0.681	210	0.808
22	115	0.489	116	0.399	116	0.678	115	0.805
23	119	0.485	119	0.398	107	0.671	116	0.805
24	121	0.472	207	0.389	115	0.653	103	0.794
25	134	0.471	108	0.388	109	0.648	207	0.791
26	203	0.470	123	0.382	112	0.641	206	0.789
27	123	0.464	203	0.372	114	0.636	123	0.786
28	207	0.461	135	0.362	101	0.636	135	0.782
29	104	0.452	103	0.361	207	0.635	114	0.781
30	103	0.450	109	0.353	123	0.635	109	0.781
31	201	0.445	134	0.351	206	0.634	118	0.780
32	135	0.441	121	0.342	108	0.634	213	0.773
33	109	0.423	201	0.329	135	0.633	209	0.767
34	122	0.422	104	0.321	102	0.620	108	0.745
35	102	0.419	112	0.316	103	0.617	102	0.717
36	112	0.412	133	0.310	201	0.579	201	0.703
37	133	0.409	122	0.297	211	0.570	132	0.694
38	211	0.397	211	0.292	120	0.529	120	0.654
39	120	0.374	206	0.282	122	0.525	211	0.648
40	206	0.363	120	0.249	214	0.515	122	0.634
41	132	0.344	132	0.244	133	0.512	133	0.619
42	214	0.329	214	0.202	132	0.491	214	0.602
43	130	0.188	130	0.120	130	0.455	130	0.516
44	129	0.120	129	0.085	129	0.260	129	0.333
45	117	0.097	117	0.035	117	0.126	117	0.191
46	212	0.072	212	0.025	212	0.064	212	0.105

Table 3-13. Comparisons of Model Rankings Based on RWFMS(5,0.5): Nominal and IDT Area-Based for a 0.01 ng m^{-3} Concentration Threshold and for IDT_{Dos} and IDT_{AP} for a 7.2 ng min m^{-3} Dosage Threshold

Rank	Model	Nominal	Model	IDT	Model	IDT_{Dos}	Model	IDT_{AP}
1	113	0.402	205	0.421	205	0.795	104	0.881
2	101	0.401	113	0.413	104	0.745	205	0.880
3	114	0.396	114	0.400	113	0.744	110	0.879
4	123	0.394	101	0.390	110	0.739	113	0.869
5	135	0.388	110	0.363	112	0.717	101	0.860
6	103	0.384	105	0.346	127	0.690	127	0.850
7	110	0.384	106	0.343	101	0.661	114	0.833
8	205	0.381	123	0.324	123	0.654	112	0.823
9	207	0.355	112	0.314	105	0.652	123	0.808
10	115	0.348	115	0.311	208	0.637	208	0.787
11	116	0.338	127	0.302	121	0.630	206	0.776
12	106	0.334	202	0.282	106	0.628	115	0.773
13	112	0.327	103	0.277	114	0.627	121	0.760
14	127	0.325	135	0.262	115	0.627	106	0.751
15	105	0.322	207	0.261	204	0.611	105	0.710
16	202	0.316	204	0.258	202	0.597	204	0.705
17	203	0.314	109	0.228	131	0.590	131	0.686
18	204	0.289	203	0.222	206	0.583	207	0.677
19	109	0.281	208	0.222	111	0.565	103	0.675
20	104	0.277	116	0.215	210	0.557	135	0.650
21	107	0.276	210	0.214	209	0.555	210	0.647
22	210	0.274	131	0.212	109	0.548	109	0.642
23	214	0.270	209	0.209	134	0.516	203	0.633
24	209	0.269	111	0.203	203	0.504	134	0.623
25	208	0.268	104	0.196	207	0.492	111	0.615
26	128	0.263	107	0.196	213	0.490	116	0.604
27	201	0.254	201	0.194	103	0.478	202	0.604
28	206	0.249	206	0.193	135	0.466	209	0.566
29	121	0.240	128	0.189	201	0.464	128	0.538
30	131	0.240	213	0.187	116	0.436	107	0.529
31	108	0.239	102	0.171	107	0.433	213	0.520
32	111	0.238	121	0.162	119	0.430	119	0.520
33	213	0.222	108	0.150	128	0.416	132	0.507
34	118	0.217	214	0.143	118	0.406	201	0.441
35	134	0.192	118	0.141	102	0.406	118	0.438
36	119	0.179	119	0.136	132	0.332	102	0.400
37	122	0.167	134	0.132	108	0.328	108	0.399
38	102	0.149	122	0.100	214	0.322	214	0.356
39	211	0.140	211	0.098	211	0.269	120	0.290
40	133	0.138	133	0.092	130	0.238	211	0.289
41	120	0.133	132	0.089	122	0.232	122	0.282
42	132	0.123	120	0.075	120	0.223	133	0.258
43	130	0.061	130	0.039	133	0.191	130	0.240
44	129	0.027	129	0.019	129	0.066	129	0.091
45	117	0.022	117	0.007	117	0.029	117	0.045
46	212	0.016	212	0.005	212	0.013	212	0.023

Table 3-14. Comparisons of Model Rankings Based on OSF: Nominal and IDT Area-Based for a 0.1 ng m^{-3} Concentration Threshold and for IDT_{Dos} and IDT_{AP} for a 72 ng min m^{-3} Dosage Threshold

Rank	Model	Nominal	Model	IDT	Model	IDT_{Dos}	Model	IDT_{AP}
1	208	0.381	208	0.442	208	0.229	105	0.145
2	128	0.411	105	0.468	105	0.241	127	0.170
3	202	0.419	202	0.474	127	0.260	208	0.171
4	101	0.424	128	0.493	134	0.287	202	0.183
5	127	0.440	127	0.501	202	0.287	128	0.197
6	107	0.446	114	0.540	128	0.292	106	0.201
7	105	0.451	101	0.544	119	0.296	119	0.217
8	131	0.462	210	0.569	210	0.317	121	0.218
9	118	0.476	209	0.573	113	0.317	134	0.219
10	115	0.481	106	0.578	111	0.328	111	0.227
11	205	0.488	113	0.580	121	0.333	101	0.233
12	134	0.492	131	0.588	209	0.342	204	0.237
13	106	0.494	115	0.595	106	0.342	113	0.237
14	210	0.495	118	0.611	114	0.356	131	0.239
15	114	0.499	107	0.613	131	0.356	112	0.244
16	111	0.505	205	0.618	213	0.379	110	0.246
17	209	0.509	213	0.618	101	0.383	132	0.246
18	204	0.513	111	0.623	205	0.387	104	0.249
19	213	0.522	119	0.635	107	0.393	107	0.250
20	110	0.526	110	0.645	115	0.397	210	0.255
21	133	0.526	207	0.658	110	0.411	205	0.256
22	119	0.547	133	0.661	204	0.413	115	0.262
23	113	0.560	204	0.666	104	0.416	209	0.275
24	207	0.562	123	0.667	207	0.430	103	0.280
25	123	0.565	134	0.687	102	0.436	213	0.283
26	102	0.572	102	0.710	118	0.436	114	0.293
27	201	0.594	201	0.717	123	0.442	123	0.295
28	203	0.606	103	0.724	112	0.450	207	0.300
29	211	0.612	203	0.736	108	0.451	206	0.308
30	121	0.637	108	0.741	135	0.455	203	0.319
31	108	0.638	121	0.754	132	0.459	109	0.330
32	104	0.647	122	0.772	103	0.467	102	0.338
33	103	0.652	109	0.774	133	0.471	211	0.350
34	122	0.653	104	0.776	116	0.474	116	0.351
35	120	0.671	135	0.783	206	0.477	201	0.352
36	135	0.687	116	0.787	203	0.480	135	0.363
37	116	0.694	211	0.799	109	0.498	118	0.364
38	109	0.695	206	0.842	211	0.503	108	0.369
39	112	0.741	120	0.851	201	0.537	133	0.376
40	206	0.778	112	0.855	122	0.566	122	0.407
41	132	0.803	214	0.951	120	0.576	120	0.432
42	214	0.817	129	0.974	214	0.596	214	0.455
43	129	0.822	132	0.980	130	0.685	130	0.634
44	117	0.927	117	1.074	129	0.747	129	0.731
45	212	1.071	212	1.110	117	0.912	117	0.829
46	130	1.092	130	1.222	212	0.963	212	0.930

Table 3-15. Comparisons of Model Rankings Based on RWFMS(1,1): Nominal and IDT Area-Based for a 0.1 ng m^{-3} Concentration Threshold and for IDT_{Dos} and IDT_{AP} for a 72 ng min m^{-3} Dosage Threshold

Rank	Model	Nominal	Model	IDT	Model	IDT_{Dos}	Model	IDT_{AP}
1	208	0.577	208	0.524	105	0.729	105	0.823
2	128	0.551	105	0.503	208	0.721	127	0.806
3	202	0.545	202	0.498	127	0.713	208	0.786
4	101	0.544	128	0.483	202	0.685	202	0.777
5	127	0.526	127	0.476	134	0.664	106	0.761
6	107	0.521	114	0.444	128	0.658	128	0.756
7	105	0.517	101	0.441	119	0.656	101	0.748
8	131	0.508	210	0.425	113	0.648	110	0.747
9	118	0.497	209	0.421	210	0.640	104	0.746
10	115	0.493	131	0.413	111	0.623	119	0.741
11	205	0.487	106	0.411	121	0.623	205	0.736
12	134	0.484	113	0.410	106	0.623	121	0.735
13	106	0.482	115	0.399	209	0.621	134	0.734
14	210	0.481	107	0.394	131	0.608	111	0.726
15	114	0.478	213	0.392	205	0.608	131	0.722
16	111	0.473	111	0.389	101	0.605	113	0.717
17	209	0.470	118	0.384	114	0.603	112	0.714
18	204	0.467	205	0.374	110	0.587	204	0.713
19	213	0.460	119	0.374	213	0.581	107	0.706
20	110	0.456	204	0.355	115	0.578	132	0.704
21	133	0.455	207	0.354	104	0.578	115	0.703
22	119	0.439	110	0.350	107	0.565	210	0.699
23	113	0.428	133	0.348	204	0.551	123	0.686
24	207	0.426	134	0.345	123	0.547	103	0.682
25	102	0.423	102	0.331	118	0.545	209	0.679
26	123	0.421	123	0.330	207	0.536	206	0.669
27	201	0.403	203	0.313	102	0.529	213	0.668
28	203	0.399	201	0.310	112	0.521	114	0.664
29	211	0.396	108	0.310	108	0.519	207	0.661
30	121	0.379	121	0.303	135	0.513	203	0.637
31	108	0.378	122	0.292	132	0.510	109	0.627
32	122	0.368	103	0.290	103	0.506	118	0.620
33	104	0.365	135	0.285	206	0.505	211	0.617
34	120	0.354	116	0.285	133	0.504	102	0.616
35	103	0.351	109	0.280	116	0.498	116	0.605
36	135	0.345	211	0.277	203	0.493	108	0.602
37	116	0.341	104	0.265	109	0.478	201	0.601
38	109	0.336	120	0.236	211	0.476	133	0.592
39	112	0.306	112	0.231	201	0.447	135	0.592
40	132	0.270	206	0.216	122	0.427	122	0.558
41	206	0.269	214	0.186	120	0.417	120	0.543
42	214	0.263	132	0.178	214	0.406	214	0.513
43	129	0.190	129	0.093	130	0.345	130	0.381
44	130	0.129	130	0.073	129	0.251	129	0.267
45	117	0.125	117	0.030	117	0.100	117	0.170
46	212	0.060	212	0.014	212	0.037	212	0.070

Table 3-16. Comparisons of Model Rankings Based on RWFMS(5,0.5): Nominal and IDT Area-Based for a 0.1 ng m^{-3} Concentration Threshold and for IDT_{Dos} and IDT_{AP} for a 72 ng min m^{-3} Dosage Threshold

Rank	Model	Nominal	Model	IDT	Model	IDT_{Dos}	Model	IDT_{AP}
1	101	0.438	101	0.432	205	0.723	104	0.828
2	110	0.409	205	0.379	110	0.720	110	0.817
3	123	0.385	105	0.373	104	0.685	205	0.799
4	205	0.381	110	0.361	127	0.673	127	0.762
5	208	0.380	123	0.358	101	0.671	101	0.748
6	202	0.377	202	0.350	105	0.669	105	0.719
7	115	0.357	106	0.336	202	0.641	123	0.692
8	106	0.357	114	0.336	123	0.625	206	0.659
9	105	0.350	113	0.326	113	0.558	106	0.637
10	128	0.344	127	0.316	115	0.552	202	0.634
11	127	0.339	115	0.311	106	0.524	115	0.615
12	207	0.323	208	0.293	206	0.517	208	0.605
13	114	0.309	103	0.268	131	0.509	131	0.603
14	113	0.305	209	0.255	209	0.506	103	0.579
15	209	0.301	128	0.251	210	0.491	112	0.574
16	107	0.300	207	0.248	208	0.486	121	0.548
17	103	0.294	210	0.240	121	0.456	113	0.547
18	210	0.282	201	0.223	114	0.448	207	0.542
19	201	0.275	107	0.211	112	0.424	210	0.523
20	204	0.271	204	0.210	103	0.421	114	0.518
21	131	0.265	104	0.204	213	0.419	128	0.509
22	134	0.246	131	0.203	128	0.404	209	0.504
23	104	0.238	111	0.184	204	0.393	204	0.498
24	213	0.227	213	0.180	134	0.389	132	0.492
25	111	0.221	206	0.180	207	0.387	109	0.463
26	203	0.217	109	0.178	111	0.368	134	0.462
27	118	0.207	203	0.165	119	0.367	213	0.453
28	109	0.198	102	0.159	109	0.345	119	0.447
29	102	0.196	121	0.153	102	0.341	111	0.444
30	206	0.193	135	0.145	107	0.317	107	0.402
31	211	0.188	112	0.145	201	0.308	135	0.365
32	121	0.185	134	0.144	132	0.307	201	0.360
33	133	0.183	119	0.141	135	0.288	102	0.334
34	135	0.182	118	0.131	203	0.266	203	0.332
35	112	0.180	116	0.125	211	0.244	116	0.311
36	119	0.176	108	0.122	116	0.228	214	0.285
37	122	0.166	133	0.117	214	0.228	211	0.282
38	108	0.166	122	0.117	108	0.220	108	0.262
39	116	0.154	211	0.115	118	0.208	133	0.260
40	120	0.145	214	0.104	133	0.195	118	0.259
41	214	0.140	120	0.078	130	0.187	122	0.251
42	132	0.099	132	0.063	122	0.186	120	0.215
43	130	0.051	130	0.027	120	0.160	130	0.170
44	129	0.047	129	0.021	129	0.064	129	0.070
45	117	0.030	117	0.006	117	0.022	117	0.040
46	212	0.014	212	0.003	212	0.008	212	0.015

Table 3-17. Comparisons of Model Rankings Based on OSF: Nominal and IDT Area-Based for a 0.5 ng m^{-3} Concentration Threshold and for IDT_{Dos} and IDT_{AP} for a 360 ng min m^{-3} Dosage Threshold

Rank	Model	Nominal	Model	IDT	Model	IDT_{Dos}	Model	IDT_{AP}
1	127	0.600	127	0.724	127	0.497	127	0.402
2	107	0.632	107	0.724	105	0.558	131	0.431
3	134	0.635	128	0.728	107	0.575	105	0.434
4	118	0.646	111	0.756	128	0.595	119	0.442
5	128	0.658	134	0.780	208	0.596	202	0.477
6	208	0.676	208	0.784	134	0.603	107	0.477
7	111	0.676	105	0.793	202	0.610	106	0.482
8	133	0.682	106	0.818	209	0.610	128	0.494
9	131	0.687	205	0.824	113	0.614	113	0.502
10	205	0.704	110	0.842	205	0.620	112	0.516
11	209	0.709	209	0.851	131	0.621	121	0.518
12	105	0.714	133	0.858	106	0.624	101	0.524
13	119	0.724	118	0.859	111	0.633	209	0.528
14	110	0.735	202	0.882	110	0.664	208	0.537
15	101	0.746	210	0.918	121	0.666	205	0.539
16	210	0.750	131	0.929	203	0.693	110	0.540
17	202	0.750	113	0.937	101	0.693	201	0.540
18	113	0.752	213	0.944	119	0.710	133	0.541
19	106	0.756	101	0.945	204	0.725	111	0.544
20	213	0.766	207	0.946	201	0.726	134	0.548
21	123	0.804	204	0.951	112	0.732	203	0.553
22	207	0.807	201	0.953	133	0.738	123	0.554
23	121	0.822	123	0.959	123	0.739	104	0.558
24	204	0.832	119	0.967	103	0.744	103	0.562
25	203	0.841	121	0.982	109	0.746	109	0.578
26	201	0.846	203	1.005	104	0.761	206	0.588
27	102	0.870	104	1.019	210	0.767	204	0.605
28	104	0.874	211	1.021	108	0.773	115	0.607
29	115	0.876	115	1.029	207	0.779	207	0.609
30	116	0.879	120	1.042	213	0.797	102	0.610
31	211	0.920	103	1.062	115	0.800	120	0.612
32	120	0.922	108	1.086	206	0.801	122	0.617
33	122	0.928	206	1.088	118	0.818	118	0.628
34	132	0.948	102	1.102	122	0.830	211	0.658
35	103	0.954	116	1.103	102	0.838	108	0.661
36	108	0.977	122	1.125	120	0.846	213	0.683
37	206	0.993	114	1.167	211	0.849	116	0.694
38	112	0.995	112	1.193	129	0.968	214	0.700
39	114	1.002	135	1.211	214	0.981	210	0.720
40	129	1.025	132	1.214	116	0.987	135	0.745
41	214	1.028	214	1.225	135	1.020	132	0.771
42	135	1.052	129	1.242	132	1.030	117	0.931
43	117	1.122	212	1.271	114	1.047	129	0.983
44	109	1.156	109	1.287	212	1.063	114	0.994
45	212	1.173	117	1.312	117	1.070	212	1.088
46	130	1.210	130	1.358	130	1.179	130	1.100

Table 3-18. Comparisons of Model Rankings Based on RWFMS(1,1): Nominal and IDT Area-Based for a 0.5 ng m^{-3} Concentration Threshold and for IDT_{Dos} and IDT_{AP} for a 360 ng min m^{-3} Dosage Threshold

Rank	Model	Nominal	Model	IDT	Model	IDT_{Dos}	Model	IDT_{AP}
1	127	0.401	107	0.313	127	0.480	127	0.564
2	107	0.381	127	0.310	105	0.431	105	0.548
3	134	0.371	111	0.303	107	0.417	131	0.535
4	118	0.368	128	0.277	128	0.396	119	0.524
5	111	0.352	134	0.263	208	0.395	106	0.505
6	133	0.349	208	0.252	134	0.394	202	0.505
7	128	0.344	118	0.237	209	0.388	107	0.495
8	131	0.341	133	0.237	131	0.387	128	0.482
9	208	0.338	209	0.226	202	0.381	113	0.477
10	209	0.325	205	0.226	111	0.381	112	0.465
11	119	0.323	105	0.224	113	0.380	121	0.464
12	205	0.319	210	0.203	205	0.373	101	0.457
13	101	0.309	131	0.195	106	0.369	209	0.455
14	210	0.305	106	0.193	121	0.359	110	0.452
15	113	0.300	101	0.192	203	0.340	205	0.447
16	105	0.299	119	0.187	110	0.334	208	0.446
17	213	0.290	204	0.187	101	0.330	133	0.446
18	110	0.278	213	0.186	119	0.328	201	0.444
19	202	0.272	110	0.184	204	0.313	111	0.444
20	106	0.267	113	0.178	133	0.313	134	0.439
21	121	0.265	121	0.178	112	0.304	203	0.436
22	207	0.264	202	0.169	109	0.303	123	0.431
23	123	0.256	203	0.168	201	0.297	104	0.430
24	204	0.255	207	0.161	210	0.294	103	0.424
25	203	0.253	115	0.151	108	0.291	109	0.416
26	115	0.235	201	0.146	213	0.278	206	0.401
27	116	0.232	120	0.145	207	0.277	204	0.400
28	201	0.231	123	0.143	123	0.274	115	0.397
29	102	0.230	211	0.142	115	0.271	102	0.397
30	104	0.212	108	0.131	103	0.267	120	0.393
31	120	0.207	116	0.120	104	0.261	122	0.392
32	211	0.201	104	0.108	122	0.252	207	0.390
33	122	0.198	103	0.106	120	0.251	118	0.372
34	132	0.196	102	0.106	102	0.248	211	0.364
35	108	0.183	122	0.099	211	0.244	108	0.359
36	103	0.180	114	0.095	118	0.232	213	0.347
37	112	0.170	112	0.084	206	0.227	116	0.341
38	114	0.168	206	0.081	116	0.178	214	0.336
39	214	0.157	135	0.076	214	0.178	210	0.325
40	129	0.155	132	0.074	129	0.174	135	0.309
41	206	0.148	214	0.069	135	0.162	132	0.254
42	135	0.143	129	0.059	114	0.148	114	0.173
43	117	0.103	109	0.047	132	0.109	129	0.151
44	109	0.099	130	0.018	130	0.087	130	0.125
45	130	0.072	117	0.016	117	0.046	117	0.096
46	212	0.059	212	0.009	212	0.027	212	0.041

Table 3-19. Comparisons of Model Rankings Based on RWFMS(5,0.5): Nominal and IDT Area-Based for a 0.5 ng m^{-3} Concentration Threshold and for IDT_{Dos} and IDT_{AP} for a 360 ng min m^{-3} Dosage Threshold

Rank	Model	Nominal	Model	IDT	Model	IDT_{Dos}	Model	IDT_{AP}
1	128	0.301	128	0.303	105	0.488	105	0.574
2	110	0.273	105	0.246	110	0.479	106	0.557
3	105	0.270	106	0.244	106	0.458	110	0.543
4	134	0.261	208	0.213	205	0.456	202	0.511
5	208	0.256	110	0.213	202	0.427	104	0.476
6	127	0.246	127	0.207	127	0.384	205	0.458
7	202	0.245	134	0.202	128	0.349	127	0.451
8	106	0.242	107	0.198	113	0.341	113	0.422
9	205	0.235	205	0.191	208	0.339	103	0.395
10	107	0.211	202	0.174	209	0.289	123	0.391
11	131	0.200	209	0.156	134	0.289	206	0.377
12	209	0.198	111	0.133	107	0.287	131	0.373
13	133	0.177	133	0.132	103	0.275	128	0.370
14	123	0.175	201	0.128	123	0.268	208	0.353
15	213	0.169	207	0.126	104	0.245	201	0.318
16	113	0.169	123	0.126	201	0.241	101	0.312
17	210	0.159	113	0.123	131	0.227	209	0.307
18	201	0.156	210	0.115	101	0.226	112	0.302
19	207	0.155	131	0.114	206	0.225	107	0.296
20	111	0.153	213	0.109	112	0.204	134	0.286
21	104	0.147	204	0.103	204	0.195	207	0.279
22	101	0.146	101	0.102	121	0.187	109	0.277
23	118	0.143	104	0.100	203	0.183	121	0.263
24	119	0.142	211	0.090	207	0.174	133	0.262
25	204	0.132	121	0.085	111	0.173	203	0.261
26	102	0.128	118	0.083	109	0.173	119	0.257
27	211	0.115	119	0.083	133	0.160	115	0.228
28	121	0.114	103	0.083	210	0.155	204	0.214
29	122	0.111	115	0.077	115	0.150	122	0.211
30	103	0.107	203	0.076	122	0.144	111	0.208
31	120	0.103	206	0.075	102	0.140	102	0.207
32	206	0.103	120	0.074	213	0.134	214	0.175
33	115	0.102	102	0.067	211	0.132	116	0.168
34	203	0.102	122	0.059	119	0.129	120	0.162
35	116	0.090	116	0.057	108	0.116	210	0.157
36	132	0.073	108	0.052	120	0.109	211	0.156
37	108	0.071	114	0.040	214	0.085	213	0.147
38	114	0.060	112	0.036	116	0.073	135	0.146
39	112	0.060	135	0.033	135	0.068	108	0.142
40	214	0.057	214	0.032	118	0.067	118	0.128
41	129	0.052	132	0.024	114	0.064	132	0.072
42	135	0.049	109	0.018	129	0.055	114	0.066
43	130	0.037	129	0.018	130	0.041	130	0.048
44	109	0.034	130	0.009	132	0.028	129	0.043
45	117	0.030	117	0.004	117	0.010	117	0.021
46	212	0.014	212	0.002	212	0.006	212	0.009

Table 3-20. Median / Mean Absolute Ranking Difference Between Various MOE Value (Nominal, IDT, IDT_{Dos}, and IDT_{AP}) Rankings for 3 Thresholds and for 3 Scoring Functions

Threshold (ng min m^{-3}), Scoring Function	Nominal - IDT	Nominal - IDT_{Dos}	Nominal - IDT_{AP}	IDT - IDT_{Dos}	IDT - IDT_{AP}	IDT_{Dos} - IDT_{AP}
7.2, OSF	2 / 3.0	4 / 6.0	5 / 6.9	4 / 7.0	5 / 7.2	3 / 5.2
7.2, RWFMS(1,1)	1.5 / 2.7	3 / 5.4	4 / 6.9	5 / 6.3	4 / 7.2	4 / 5.1
7.2, RWFMS(5,0.5)	2 / 3.2	5.5 / 7.0	4 / 6.2	4 / 5.1	4 / 4.8	1 / 2.6
72, OSF	2 / 3.1	5 / 5.9	6 / 7.7	3 / 5.1	4 / 7.5	4 / 4.5
72, RWFMS(1,1)	2 / 3.0	5 / 6.0	7 / 7.5	3.5 / 5.3	6 / 7.7	4 / 4.7
72, RWFMS(5,0.5)	2 / 2.6	3.5 / 5.7	4 / 6.2	3 / 4.9	3 / 5.2	1 / 2.2
360, OSF	2 / 2.9	5 / 6.0	5 / 7.4	4 / 5.1	5.5 / 7.3	3 / 4.1
360, RWFMS(1,1)	2 / 2.7	4 / 5.9	5 / 7.7	3 / 5.0	6 / 8.2	3.5 / 4.8
360, RWFMS(5,0.5)	2 / 2.4	5 / 5.9	5 / 7.4	4 / 5.3	6 / 7.3	2.5 / 3.3

absolute rankings for the 46 models for each of the six possible ranking comparisons.

The results that can be gleaned from Table 3-20 are quite similar to those that were seen for the VW-based comparisons (Table 3-10). Again, differences in model rankings are greatest when comparing concentration-based to dosage-based MOE values. And also again, some models (e.g., SCIPUFF) relative rankings improve greatly when dosage is considered instead of concentration.

C. SUMMARY

Table 3-21 provides a summary of how model rankings change as a function of MOE type when focusing on the two area-based procedures – VW_c and IDT. The values reported in Table 3-21 correspond to the median difference in red and the mean difference in blue for the absolute rankings for the 46 models. Three "paired" comparisons are considered: VW_c versus IDT, VW_{cDos} versus IDT_{Dos}, and VW_{cAP} versus IDT_{AP}. This table along with Tables 3-10 and 3-20 suggest that the biggest changes in relative model rankings are due to considering dosages instead of concentrations (as previously discussed). The next biggest changes in relative model rankings appear to be due to differences in the area-based technique adopted – VW_c or IDT. Finally, inclusion of actual European population distributions vice a uniform population (the "$_{Dos}$" cases), leads to the smallest changes in relative rankings, at least for the thresholds that we examined.

Table 3-21. Median / Mean Absolute Ranking Difference Between Various MOE Value (VW_c and IDT) Rankings for 3 Thresholds and for 3 Scoring Functions

Threshold (ng min m^{-3}), Scoring Function	VW_c - IDT	VW_{cDos} – IDT_{Dos}	VW_{cAP} – IDT_{AP}
7.2, OSF	3 / 5.0	5 / 6.3	5 / 6.1
7.2, RWFMS(1,1)	2 / 4.7	5 / 6.5	5 / 6.2
7.2, RWFMS(5,0.5)	3 / 5.3	5 / 6.2	5.5 / 6.0
72, OSF	6 / 7.2	5 / 6.7	5 / 6.4
72, RWFMS(1,1)	5 / 6.7	4.5 / 6.5	4 / 5.5
72, RWFMS(5,0.5)	3.5 / 6.0	3.5 / 5.0	4 / 5.3
360, OSF	4 / 4.8	5 / 6.0	6.5 / 9.4
360, RWFMS(1,1)	5 / 6.8	4 / 5.6	7 / 9.3
360, RWFMS(5,0.5)	3 / 3.5	3.5 / 4.8	5 / 6.2

Table 3-22 lists the models that were in the top 5 (or top 10) of 46 under more than one computational procedure. This table describes results for the three scoring functions and three threshold levels (simply labeled low, medium, and high since both concentration and dosage thresholds were considered here). The second column identifies those model predictions that were always in the top 5 (top 10) for all techniques (labeled "All"): nominal, VW_{80}, VW_c, IDT, VW_{cDos}, IDT_{Dos}, VW_{cAP}, and IDT_{AP}. The third column identifies those model predictions that were always in the top 5 (top 10) for all techniques that involved some type of accounting for the area (labeled "All-Nominal"): VW_{80}, VW_c, IDT, VW_{cDos}, IDT_{Dos}, VW_{cAP}, and IDT_{AP}. The final column identifies those model predictions that were always in the top 5 (top 10) for all techniques that were dosage-based (labeled "Dosage/Population-Based"): VW_{cDos}, IDT_{Dos}, VW_{cAP}, and IDT_{AP}. Therefore, Table 3-22 illustrates the models that had robust performance (top 5 or top 10) with respect to the MOE value computational technique.

Table 3-22. Robust Top 5 / Top 10 Ranked Models for 3 Thresholds and for 3 Scoring Functions. For thresholds, "Low" implies 0.01 ng m^{-3} and 7.2 ng min m^{-3} for concentration and dosage measures, respectively; "Medium" implies 0.1 ng m^{-3} and 72 ng min m^{-3} for concentration and dosage measures, respectively; and "High" implies 0.5 ng m^{-3} and 360 ng min m^{-3} for concentration and dosage measures, respectively.

Threshold, Scoring Function	All	All – Nominal	Dosage/Population-Based
Low, OSF	208 / 105	208 / 105	208 / 105, 111, 121, 131
Low, RWFMS(1,1)	208 / 105	208 / 105	208 / 105, 121, 131
Low, RWFMS(5,0.5)	- / 123, 205	- / 123, 205	104 / 123, 205
Medium, OSF	- / 105, 127, 202, 208	- / 105, 127, 202, 208	105, 127, 208 / 119, 202
Medium, RWFMS(1,1)	- / 105, 127, 202, 208	- / 105, 127, 202, 208	105, 127, 208 / 119, 202
Medium, RWFMS(5,0.5)	- / 101, 105	- / 101, 105, 127	104, 110, 127, 205 / 101, 105
High, OSF	127 / 107, 128	127 / 107, 128	127 / 105, 107, 128
High, RWFMS(1,1)	127 / 107	127 / 107	127 / 107, 128
High, RWFMS(5,0.5)	110 / 105, 127, 202, 205	110 / 105, 127, 202, 205	105, 110 / 106, 127, 202, 205

For the lowest thresholds considered, models 208 (SMHI) and 105 (CMC) have the most robust "top 5 / top 10" performance for the OSF and RWFMS(1,1) scoring functions. For the more conservative RWFMS(5, 0.5) scoring function and the lowest thresholds, models 123 (Meteo) and 205 (DMI) resulted in the most robust top 5 / top 10 performance. For the medium thresholds, models 105 / 202 (CMC), 127 (ARAC), and 208 (SMHI) showed the most robust top 5 / top 10 performance for the OSF and RWFMS(1,1) scoring functions. For the more conservative RWFMS(5, 0.5) scoring function and the medium threshold, models 101 (IMP) and 105 (CMC) resulted in the most robust top 5 / top 10 performance. For the highest thresholds considered, models 107 (DWD), 127 (ARAC), and 128 (SMHI) displayed the most robust top 5 / top 10 performance for the OSF and RWFMS(1,1) scoring functions. For the more conservative RWFMS(5, 0.5) scoring function and the highest threshold, models 105 / 202 (CMC), 110 / 205 (DMI), and 127 (ARAC) resulted in the most robust top 5 / top 10 performance. The CMC (Canadian Meteorological Centre, model 105) predictions appear in the top 5 / top 10 often *and* for both the OSF/RWFMS(1,1) and the more conservative RWFMS(5,0.5) scoring functions. This implies an important level of robustness for these predictions. The Lawrence Livermore National Laboratory (LLNL) model (ARAC, model 127) also showed a relatively robust performance as judged by the top 5 / top 10 rankings and the nominal and conservative scoring functions reported in Table 3-22.

Finally, an important caveat must be noted. To this point, the use of area-based and population based MOE values to compare sets of model predictions, rank the models, and provide insight into relative model performance has been emphasized. With respect to the population-based MOE values, it has been previously noted that the x-axis can correspond to one minus the fraction of the (exposed) population that is inadvertently exposed to a threshold level of interest and the y-axis can correspond to one minus the fraction of the (warned) population that is unnecessarily warned (at a threshold level of interest). One might imagine using an effects (or lethality) model to compute, via minimal extension of the MOE, the actual number of people "falsely warned" or "inadvertently exposed." However, one must be careful because of the relatively small number of samplers associated with the observed *ETEX* data. Recall, 168 samplers were used to cover all of Europe. In attempting to describe the actual number of effected people, one would need to compute the absolute (actual) areas, not simply the fraction of areas. In such a case, the estimated area sizes can be sensitive to the details associated with the specific area-based technique (e.g., interpolation). For example, Figure 3-3 shows contours for the observations based on the IDT technique both with and without initial logarithmic transformation. It is seen that the area sizes greatly differ depending on this choice. Two dosage contours are shown in Figure 3-3 – 1.8 and 360 ng min m^{-3}. Table 3-23 shows the differences in area size and population associated with the two contour levels and the two IDT techniques ("linear," that is, no transformation of the observations and "log," that is, initial logarithmic transformation of all observations).

Please note that this sensitivity is solely based on the availability of experimental data. As in *ETEX*, data is often collected at sparsely distributed samplers. In general, predictions produce plumes on regular grids and do not necessarily require any interpolation. Thus, one might envision the following operational procedure to access actual "areas" or "numbers of people affected." First, models are compared to observations using data at the samplers, both observed and predicted, as was done in this study. Next, one chooses a model prediction that demonstrates robust and acceptable performance – that is, "top ten" performance no matter what the MOE computational technique. Then, the corresponding "robust model" predicted plume could be used to calculate actual areas and/or numbers of people affected. We did not have access to the predicted plumes for this study, only the predicted concentrations at the sampler locations.

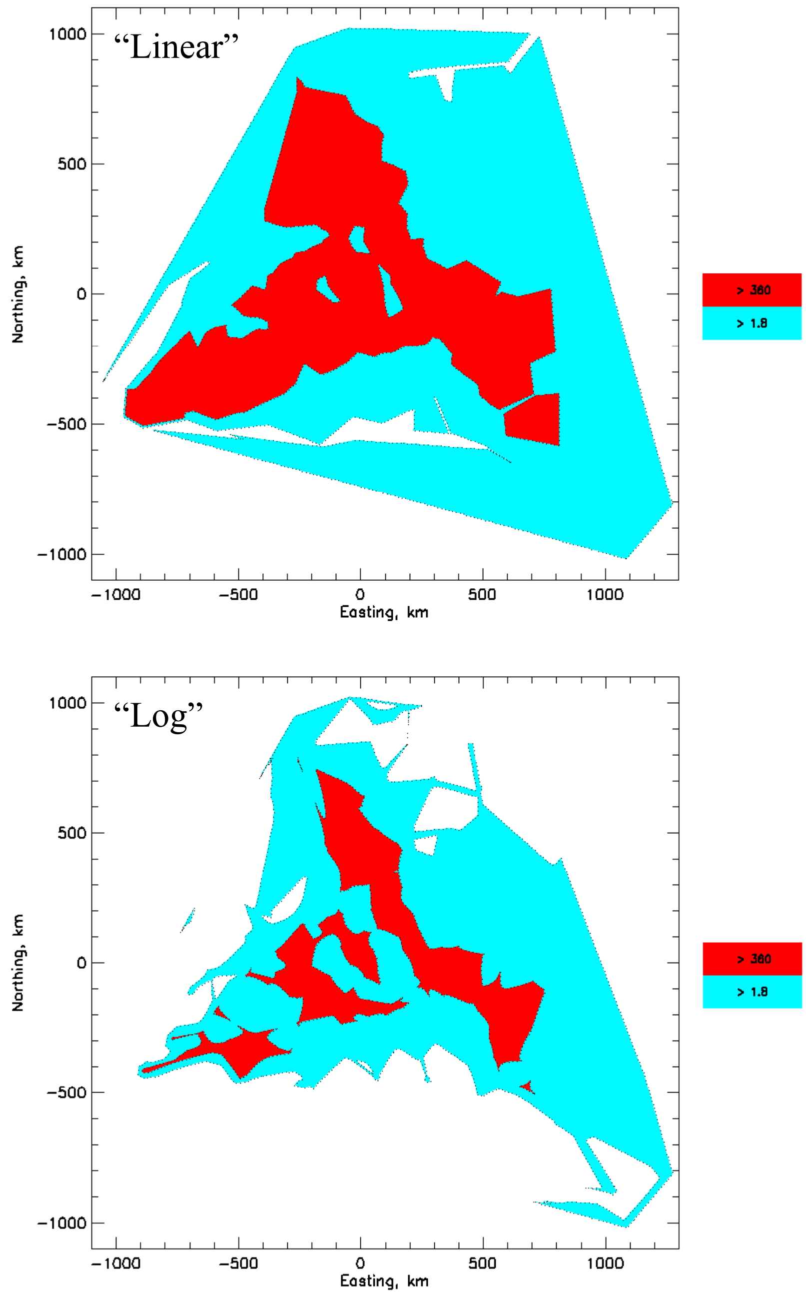


Figure 3-3. Contours (Based on IDT) for Dosage Thresholds of 1.8 and 360 ng min m^{-3} for Observed *ETEX* Data

Table 3-23. Absolute (actual) Values of Area/Population Contained Within 2 Dosage Thresholds (1.8 and 360 ng min m^{-3}). Interpolation via Delaunay Triangulation (IDT) was used to create the contours. In addition, before IDT, observed dosages are either transformed logarithmically ("Log") or left as is ("Linear").

	Area		Population	
	Contour Level		Contour Level	
	1.8	360	1.8	360
Linear	2 911 076	924 820	307 388 527	139 572 593
Log	1 777 056	373 220	220 146 360	79 934 963

REFERENCE

3-1. Warner S., Platt, N., Heagy, J. F., 2003: *Application of User-Oriented MOE to Transport and Dispersion Model Predictions of the European Tracer Experiment*, IDA Paper P-3829, 86 pp, November 2003. (Available electronically [DTIC STINET ada419433] or on CD via e-mail request to Steve Warner at swarner@ida.org or a mail request to Steve Warner, Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, Virginia 22311-1882) and Warner S., Platt, N., Heagy, J. F., 2003: Application of user-oriented MOE to transport and dispersion model predictions of the European tracer experiment. *Atmospheric Environment*, in press.

APPENDIX A

ACRONYMS

APPENDIX A
ACRONYMS

ABS	Absolute value
ABS(FB)	Absolute Fractional Bias
A_{OB}	Region Associated With the Observations
A_{FP}	False Positive Region
A_{FN}	False Negative Region
ANPA	National Agency for Environment (Italy)
A_{OV}	Region of Overlap
A_{PR}	Region Associated With the Prediction
ARA	Applied Research Associates
ARAC	Atmospheric Release Advisory Center
ARAP	Aeronautical Research Associates of Princeton
ATMES	Atmospheric Transport Model Evaluation Study
ATP	Allied Tactical Publication
BMRC	Bureau of Meteorology Research Center (Australia)
CD	Compact disc
C_{FN}	false negative coefficient
C_{FP}	false positive coefficient
CMC	Canadian Meteorology Centre
CNR	National Research Council (Italy)
DNMI	Norwegian Meteorological Institute
DMI	Danish Meteorological Institute
d_{OSF}	distance to (1,1) (for objective scoring function)
DTIC	Defense Technical Information Center
DTRA	Defense Threat Reduction Agency
DWD	German Weather Service
ECMWF	European Centre for Medium Range Weather Forecasts
EDF	France Electricity
ETEX	European Tracer Experiment
FB	Fractional Bias
FOA	Defense Research Establishment (Sweden)
FOM	Figure-Of-Merit

HPAC	Hazard Prediction and Assessment Capability
IDA	Institute for Defense Analyses
IDL	Interactive Data Language
IDT	Interpolation via Delaunay Triangulation
IDT_{AP}	Interpolation via Delaunay Triangulation Based on Actual Population
IDT_{Dos}	Interpolation via Delaunay Triangulation Based on Dosages
IMP	Institute for Meteorology and Physics, University of Wien (Austria)
IMS	Swiss Meteorological Institute
IPSN	French Institute for Nuclear Protection and Safety
JAERI	Japan Atomic Research Institute
JMA	Japan Meteorological Agency
KMI	Royal Institute of Meteorology of Belgium
LLNL	Lawrence Livermore National Laboratory
Log	Logarithm
Meteo	Meteo France
MetOff	Meteorological Office (United Kingdom)
MOE	Measure of Effectiveness
MRI	Meteorological Research Institute (Japan)
MSC-E	Meteorological Synthesizing Centre – East (Russia)
NARAC	National Atmospheric Release Advisory Center
NERI	National Environment Research Institute / Risoe National Laboratory/ University of Cologne (Germany / Denmark)
ng m^{-3}	nanograms per cubic meter
ng min m^{-3}	nanogram minutes per cubic meter
NIMH-BG	National Institute of Meteorology and Hydrology (Bulgaria)
NIMH-BG	National Institute of Meteorology and Hydrology (Romania)
NOAA	National Oceanic and Atmospheric Administration
OLAD	Over-Land Along-Wind Dispersion
OSF	Objective Scoring Function
PMCH	Perfluoro-methyl-cylcohexane
Ref.	Reference
RWFMS	Risk-Weighted Figure of Merit in Space
SAIC	Science Applications International Corporation

SCIPUFF	Second-Order Closure Integrated Puff
SMHI	Swedish Meteorological and Hydrological Institute
SRS	Westinghouse Savannah River Laboratory
T&D	Transport and Dispersion
UTC	Universal Time Coordinated
V&V	Verification and Validation
VV&A	Verification, Validation, and Accreditation
VW	Voronoi Weighted
VW_{80}	Voronoi Weighted with 80th Percentile Clipping
VW_c	Voronoi Weighted with Clipping
VW_{cAP}	Voronoi Weighted with Clipping Based on Actual Population
VW_{cDos}	Voronoi Weighted with Clipping Based on Dosages
WGS 84	World Geodetic System 1984

APPENDIX B

TRANSFORMATION BETWEEN WORLD GEODETIC SYSTEM 1984 (WGS 84) COORDINATES AND PSEUDO-RECTANGULAR COORDINATES

APPENDIX B
TRANSFORMATION BETWEEN WORLD GEODETIC SYSTEM 1984 (WGS 84) COORDINATES AND PSEUDO-RECTANGULAR COORDINATES

The WGS 84 system [Ref. B-1] represents the large-scale shape of the Earth as an ellipsoid of revolution, formed from the ellipse shown in Figure B-1.

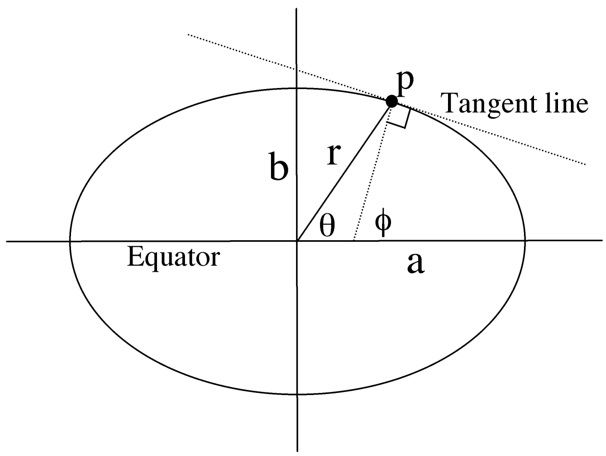


Figure B-1. Fundamental WGS 84 Ellipse.

The semi-major axis is given by a = 6378.137 km. The semi-minor axis b, is defined through the flattening parameter, given by $f = (a\text{-}b)/a$ = 1/298.257223563 = 0.0033528. The angle ϕ is known as the geodetic latitude, while θ is called the geocentric latitude. It can be shown that ϕ and θ are related through the expression

$$\sin\theta = \frac{\sin\phi\sqrt{1+A}}{\sqrt{1+A\sin^2\phi}} \tag{B-1}$$

where $A=(2f-f^2)(2f-f^2-2)$. Consider a point p_0 on the Earth with WGS 84 latitude and longitude values ϕ_0 and α_0, respectively.

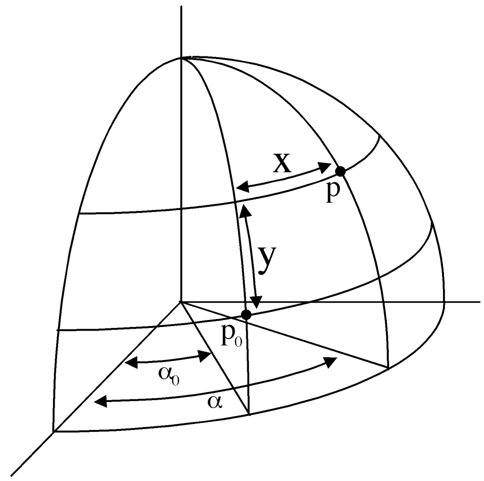


Figure B-2. WGS 84 Ellipsoid with Points p_0 and p; p_0 is the Origin of the Pseudo-Rectangular Coordinate System.

One can construct a pseudo-rectangular coordinate system with p_0 at the origin, as follows. An arbitrary point p (see Figure B-2) with WGS 84 latitude and longitude values ϕ and α has a y coordinate equal to the arc length along the meridian passing through p_0 from ϕ_0 to ϕ. The x coordinate is the arc length along the parallel passing through p from α_0 to α. The two arc lengths are found as follows.

In each case, the geometry simplifies considerably if the geodetic latitudes ϕ_0 and ϕ are converted through Eq. B-1 to the geocentric latitudes θ_0 and θ. Once this has been carried out, the x coordinate is given by

$$x = r\cos\theta(\alpha - \alpha_0). \tag{B-2}$$

From the properties of ellipses, the radius r is given by

$$r = \frac{b}{\sqrt{1-\varepsilon^2\cos^2\theta}}, \tag{B-3}$$

where $\varepsilon^2 = 1 - \frac{b^2}{a^2} = 2f - f^2 = 0.0066719$. Combining Eqs. B-2 and B-3, one has

$$x = \frac{b\cos\theta(\alpha-\alpha_0)}{\sqrt{1-\varepsilon^2\cos^2\theta}}. \tag{B-4}$$

The arc length y requires more work. An element of arc length ds along the meridian is given through the fundamental arc length relation in polar coordinates:

$$ds = \sqrt{r^2 + (\tfrac{dr}{d\theta})^2}\, d\theta. \tag{B-5}$$

Carrying out the derivative using Eq. B-3, and simplifying one has, after some algebra,

$$r^2 + (\tfrac{dr}{d\theta})^2 = \frac{b^2(1+\cos^2\theta(\varepsilon^4 - 2\varepsilon^2))}{(1-\varepsilon^2\cos^2\theta)^3}\,. \tag{B-6}$$

The arc length y is then given by the integral

$$y = \int_{\theta_0}^{\theta} ds = \int_{\theta_0}^{\theta} \left[\tfrac{b^2(1+\cos^2\xi(\varepsilon^4-2\varepsilon^2))}{(1-\varepsilon^2\cos^2\xi)^3}\right]^{\frac{1}{2}} d\xi\,. \tag{B-7}$$

This integral cannot be carried out analytically, but can be done numerically. For the precision sufficient for the MOE analyses in this work, one can expand the integrand in Eq. B-7 to order ε^2 and compute the resulting integral exactly, as follows.

To order ε^2, the integral is given by

$$y \approx \int_{\theta_0}^{\theta} b(1+\tfrac{\varepsilon^2}{2}\cos^2\xi)\,d\xi = b\Big[(\theta-\theta_0)(1+\tfrac{\varepsilon^2}{4}) + \tfrac{\varepsilon^2}{8}(\sin 2\theta - \sin 2\theta_0)\Big]. \tag{B-8}$$

The difference between Eq. B-8 and the exact arc length (Eq. B-7) is a maximum for $\theta_0 = 0$, that is, at the equator. As an example, setting the angle θ to 200/6378 = 0.0313571 (corresponding to an arc length of roughly 200 km), the difference is 3.4 meters (Eq. B-7 - Eq. B-8).

In summary, to generate pseudo-rectangular coordinates from a set of WGS 84 latitudes and longitudes, one first chooses a latitude and longitude (θ_0, ϕ_0) to serve as the origin. One then converts all latitudes to geocentric latitudes via Eq. B-1. Finally, one gets the (x, y) pairs from Eqs. B-4 and B-8, respectively.

REFERENCE

B-1. *WGS 84 Implementation Manual*, *Version 2.4,* Prepared by EUROCONTROL, European Organization for the Safety of Air Navigation Brussels, Belgium and IfEN, Institute of Geodesy and Navigation (IfEN) University FAF Munich, Germany, Feb. 1998. Available at http://www.wgs84.com/wgs84/downloads.htm.

APPENDIX C

TASK ORDER EXTRACT

APPENDIX C
TASK ORDER EXTRACT

DC-9-1797

TITLE: Support for DTRA in the Validation Analysis of Hazardous Material Transport and Dispersion Prediction Models

This task order is for work to be performed by the Institute for Defense Analyses (IDA) under Solicitation Number RFP#DASW01-04-0003 for the Defense Threat Reduction Agency (DTRA).

1. <u>BACKGROUND:</u>

The Hazard Prediction and Assessment Capability (HPAC) is a suite of codes that predicts the effects of hazardous material releases into the atmosphere and their impact on civilian and military populations. The software can use integrated source terms, high-resolution weather forecasts, and particulate transport models to predict hazard areas produced by battlefield or terrorist use of weapons of mass destruction (WMD), by conventional counterforce attacks against WMD facilities, or by military and industrial accidents.

The DTRA Verification and Validation (V&V) Program represents ongoing activities performed in parallel with development of all predictive codes in support of HPAC. One element of V&V is to perform code-on-code comparisons. In this strategy, each code receives the same input. In this manner, differences in the output predictions can lead to the identification of software bugs, or help to assess technical strengths and weaknesses of component algorithms within each code. In addition, a certain amount of credibility for both models is achieved when their predictions agree. When the inputs are simple, such as for fixed winds and simple terrain, the predictions tend to be dominated by the dispersion algorithms. Comparisons at this level of complexity are important to establish fundamental dispersion algorithm veracity, and to help discover software bugs. As more complex terrain and weather is included as input, the number of physical processes responsible for transport and dispersion increases and the predictions become the result of many interdependent algorithm calculations.

It is very difficult to separate meteorological uncertainty from the transport and dispersion model accuracy when comparing predictions to field-trial validation quality or real-world data. The validation challenge is to assess whether a model performs well over different field trials, and ultimately reflects real-world phenomena. Some codes perform better under certain conditions and specific scenarios. Hazard prediction models are generally developed for a range of user communities and applications. Each user community has a different set of requirements. Thus, the corresponding hazard models tend to be optimized for specific applications. The process of validating a model should be couched in terms of end-user requirements where feasible.

Various figures-of-merit (FOM) are used to express model performance relative to observed data. Most FOMs tend to use manifestations of a ratio (geometric or arithmetic) between the predicted and observed quantities. The compared quantities are usually peak, plume-centerline, and off-axis concentration or dosage, as well as crosswind and along-wind spread and area coverage. Other FOMs may include the second-moment of the dosage and concentration values at a sampler location. All these FOMs are reasonable measures, but none of them explicitly expresses application-oriented performance. A "yardstick" is needed that measures application-oriented model performance. The scale on this yardstick would clearly and directly relate to the specific user's concerns and needs. The pursuit of this "validation" performance measure (Measure of Effectiveness or MOE) is a continuing initiative at DTRA.

2. OBJECTIVE:

IDA will conduct independent analysis and special studies associated with verification and validation of the suite of models associated with the Hazard Assessment and Prediction Capability. IDA will support development of user-oriented performance measures of effectiveness (MOE) using validation quality field trial data sets; coordinate scenario definition and arbitration for code-on-code V&V activities.

The objectives of verification and validation analysis and coordination are: (1) to ensure that a consistent analysis approach is used when comparing model predictions, and assist DTRA in the implementation of code-on-code analysis, comparisons, and interpretation; and (2) to define and further develop measures of effectiveness in terms of user-specific objectives and applications.

The scope of this effort may be expanded to other programs as directed by DTRA.

3. STATEMENT OF WORK:

As required by DTRA technical representatives, IDA will perform the following tasks:

a. Advanced User-Oriented Measure of Effectiveness (MOE) Development

IDA will conduct model prediction to field trial observation comparisons using a novel user-oriented MOE. Mean value and probabilistic prediction outputs (e.g., from HPAC) will be examined and relative performance will be described. IDA will follow-up on their FY'03 analysis and report that compared the predictions of 46 models (including SCIPUFF and ARAC) to the observations of the 1998 ETEX release. In FY'03, a wealth of information was received from the JRC, Italy (European Commission). These data included the observations of ETEX (Release 1) and the predictions of 46 models that participated in their ATMES II study. Follow-up IDA analysis will include the development of appropriate procedures for area interpolation, sensitivity analysis, and consideration of notional affected populations.

b. Communication: Using the MOE for Model Validation

IDA will focus particular effort on the communication, via various methods, of the value, usage, and technical merits of the new validation and accreditation MOE. Technical and operator review and feedback will be sought and considered. IDA will continue the development of "demonstration" validations in the context of specific user requirements. This effort will require the identification of a potential user and specific application. For this user(s) and application(s), IDA will focus on extracting a sense for what are the acceptable user requirements (i.e., risk tolerance). These requirements will differ among potential user groups (military targeting, passive CB defense, civilian first responders, military versus civilian population human effects, etc.). Similarly, previously described lethality/effects filters will be used to interpret MOE results and reviewed with potential users. The goal of the above effort is to demonstrate the "end-to-end" validation ("accreditation") of a model usage (e.g., a particular HPAC probabilistic output) for a specific application and user (i.e., agreed to/acceptable risk tolerance). The chosen application and user should correspond to an actual situation (i.e., not simply represent a notional scenario).

(1) A follow-on FY 04 study, which will be begun in FY 03, will use archived CDMs (Chemical Downwind Messages), which are the 6 hour advanced forecasts of theater weather conditions upon which the ATP-45 templates are based. We will compare these templates to HPAC predictions using real archived weather conditions. This study will also include terrain effects (as modeled by HPAC) and real population databases for the regions of study. Our results will be framed in terms of the number of people not warned (false negatives) and the number of people unnecessarily warned (false positives). Ultimately, it is hoped that this effort will address the topic of acceptable risk tolerance within the context of the military use of a T&D code.

(2) IDA will communicate, via conference papers and/or posters, working group discussions, and IDA papers, the more important applications of the MOE and any progress toward the creation of "demonstration" validations. In addition, IDA intends to create descriptions of its efforts, where appropriate (and approved by DTRA), that are suitable for publication in peer-reviewed journals. We expect that some of our analyses associated with Urban 2000 and ETEX will be suitable for FY '04.

c. Comparisons of DTRA-Identified Urban and Building Interior T&D Models

IDA will continue (efforts begun in FY 03) to extend the application of the user-oriented MOE to building interior and urban models of hazardous material transport and dispersion. IDA plans to examine data collected during the MUST field experiment and provide an analysis of Urban HPAC predictions of the MUST field experiment. Comparisons of predictions of the transport and dispersion of pollutants within a building are also planned for FY '04 (pending data availability). Initial examinations of the observations (MET and sampler) associated with the Joint Urban 2003 experiment are also planned for late FY '04 (when these data become available).

d. Joint Validation Studies with LLNL (OPTION 1)

Pending approval and available funding, IDA will conduct data analysis and comparisons of observations and predictions associated with NARAC and HPAC predictions of previous field trials. This effort will be a collaborative effort with LLNL

(DOE). Previous field trials being considered for this study include mid-range (OLAD, DIPOLE PRIDE 26), long range (ETEX), and complex terrain (ASCOT, DIABLO CANYON) releases. In order for these comparisons to be meaningful and credible, a careful, objective protocol associated with the running of the models and the comparisons of predictions and observations will be developed and described. Issues that were discovered during the Phase I HPAC -NARAC comparisons will also be addressed, where feasible, either with additional field trial comparisons ("piggy-backing") or, at least in part via technical review. In addition, the inclusion of different predictive weather inputs ("weather experts") may be considered within the framework of model validation/accreditation.

e. Using the MOE for Model Accreditation of NBC CREST (OPTION 2)

Pending approval and separate funding, IDA will apply the concept of risk tolerance and end-to-end accreditation by reviewing the Office of the Army Surgeon General's (OTSG) developed software call NBC CREST (Casualty and Resource Estimation Support Tool). This software uses HPAC to develop the plume and is to be linked to CATS in FY04. The doctrine behind CREST is AMED P-8 for casualty estimation and AMED P-7 for CONOPS. The latter is being reviewed for an updated (revised) NATO STANAG.

4. <u>CORE STATEMENT</u>:

This research is consistent with IDA's mission in that it will support specific analytical requirements of the sponsor and will assist the sponsor with planning efforts. Accomplishment of this task order requires an organization with experience in operationally oriented issues from a joint and combined perspective, which IDA, a Federally Funded Research and Development Center, is able to provide. It draws upon IDA's core competencies in Systems Evaluations and Operational Test and Evaluation. Performance of this task order will benefit from and contribute to the long-term continuity of IDA's research program.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
1-10-2004

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*
Jan. 2004 - Aug. 2004

4. TITLE AND SUBTITLE
Comparisons of Transport and Dispersion Model Predictions of the European Tracer Experiment: Area-Based and Population-Based Measures of Effectiveness

5a. CONTRACT NUMBER
DASW01-04-C-0003

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER
DC-9-1797

5f. WORK UNIT NUMBER

6. AUTHOR(S)
Steve Warner, Nathan Platt, James Heagy

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Institue for Defense Analyses
4850 Mark Center Drive
Alexandria, VA 22311-1882

8. PERFORMING ORGANIZATION REPORT NUMBER
IDA Paper P-3915

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Defense Threat Reduction Agency
6801 Telegraph Road, Room 264
Alexandria, VA 22310-3398

10. SPONSOR/MONITOR'S ACRONYM(S)
DTRA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.
Directorate for Freedom of Information and Security Review, 21 October 2004.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
In October 1994, the tracer gas perfluoro-methyl-cyclohexane (PMCH) was released over a 12-hour period from a location in northwestern France and tracked at 168 sampling locations in 17 countries across Europe (100s of kilometers). This release, known as the European Tracer Experiment (ETEX), resulted in the collection of a wealth of data. IDA has obtained the predictions of 46 transport and dispersion models from 17 countries from the Joint Research Centre, European Commission as well as the PMCH sampling data. This paper applies a previously described user-oriented measure of effectiveness (MOE) methodology to evaluate the predictions of the 46 models against the long-range ETEX observations. The paper extends previous work by computing MOE values that are based on "true" areas (e.g., in square kilometers) and on actual European population distributions. The overall objective of this paper is to document the procedures used to estimate the area-based and population-based MOE values.

15. SUBJECT TERMS
model validation; hazardous material transport and dispersion; HPAC; ETEX; measure of effectiveness

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Mr. Richard Fry
Unclassified	Unclassified	Unclassified	Unlimited	139	**19b. TELEPHONE NUMBER** *(include area code)* 703-325-6760

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18